


```

||||| ||| :|||: ||| ||| |||
501 AGAATGGAAGATTGGGAGGAGTGAGACAG.....CTGATG 541
      |||||
89 ArgLysThrThrGluGlnValArgProTyrArgArgPheArgThrPr 105
      |||||
542 GAAAGCTGAGGAAAGCAGTTCAGTCATAGTCTGGGGCAGTCAGCAC 591
      |||
105 oclupro.....AspAsnHisTyrAspPheCysLeuIlePro 117
      |||||
592 TGATCCCTCAGCATGACCATCAGTATGATTTTGCCTATGCCC 637
      |||||

```

seq_name: /SIDS2/gcgdata/geneseq/NA1999.DAT:AA23519

seq_documentation_block:

ID_AA23519 standard; DNA: 44453 BP.

AC_AA23519;

23-JUN-1999 (first entry)

Human kidney aminopeptidase P genomic DNA fragment 3.

Aminopeptidase; human; Amp; gene therapy; treatment: Amp-deficiency;
prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;
arterial stenosis; industrial protein feed; malabsorption syndrome;
proteinaceous waste degradation; additive; immunohistochemistry; ss.

Homo sapiens.

W09911799-A2.

11-MAR-1999.

02-SEP-1998; 98WO-US18426.

02-SEP-1997; 97US-0057854.

(MEDI-) MEDICAL COLLEGE GEORGIA RES INST.

Ryan JW, Sprinkle TJC, Venema RC;

WPI; 1999-205193/17.

Nucleic acid encoding human aminopeptidase P

Claim 13; Page 139-165; 201pp; English.

This invention describes the isolation of a novel human aminopeptidase P (Amp). This protein is used to produce recombinant Amp and can be used for gene therapy for treating Amp-deficiency conditions. Its fragments are used as primers and probes to identify patients with homozygous and heterozygous Amp deficiency, including prenatal diagnosis (patients defective in Amp are at risk of developing angioedema if treated with angiotensin-converting enzyme inhibitors), also as antisense inhibitors in cases of excessive Amp expression. The product of the invention is also used to identify Amp-expressing sequences in other animals and to generate transgenic animals, and comparisons of genomic sequences are used to detect mutations. Amp inhibitors are potentially useful as antihypertensive agents and to prevent or treat arterial (re)stenosis or atherosclerosis. The structure of Amp is used to design synthetic substrates, e.g. for use in Amp assays. Amp, which hydrolyzes N-terminal imido bonds, can be used to degrade industrial protein feeds to free amino acids, to degrade proteinaceous wastes, as additives in enzyme formulations used to treat malabsorption syndrome and for studying its biological role. Antibodies against Amp are used in immunohistochemical methods to study Amp distribution.

Sequence 44453 BP; 13034 A; 10110 C; 10292 G; 11017 T; 0 other;

alignment_scores:

Quality: 98.50

Ratio: 1.539

Length: 102

Gaps: 2

Percent Similarity: 62.745 Percent Identity: 26.471

alignment_block:

US-09-327-750D-35 x AA23519

Align seg 1/1 to: AA23519 from: 1 to: 44453

```

1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAspLy 17
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43713 CTACAGAAGAGAGGACACTGGCTTTTCCAGTCAGCTTGAGGAAGAG 43762

17 slsAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProH 34
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43763 GGAGGAAAAAGAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 43812

34 IsHisLeuGluGluValGluAsnLysLysProGly..GlyAsnValArg 50
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43813 TGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 43862

50 rgLysValArgLeuValProAsnPheLeuTirAlaIleProAsnArg 66
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43863 GGAAGAAGAAGAGG.....AAGTGGAGGAGGCCAGAAAAAG 43897

67 HisValAspArgAsnGluGlyGlyGluAspValGlyArgPheValGlu 83
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43898 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 43947

83 ngLysThrGluValLysArgLysThrThrGluGlnGlnValArgProTyr 100
  : : : : : : : : : : : : : : : : : : : : : : : : : : : :
43948 AGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAA 43997

100 rgArg 101
  : : : : :
43998 GCAAA 44002

```

seq_name: /SIDS2/gcgdata/geneseq/NA2001.DAT:AAH45143

seq_documentation_block:

ID_AA45143 standard; CDNA: 792 BP.

AC_AA45143;

07-SEP-2001 (first entry)

Human brain expressed X-linked protein, hBex, coding sequence.

Human; brain expressed X-linked protein; cytostatic; auditory; nontropic;
hBex; dysembryoplasia; hereditary disease; cancer; tumour; deafness;
X-chromosome-binding mental retardation; lissencephalous disease; ss.

Homo sapiens

W0200140286-A1.

07-JUN-2001.

27-NOV-2000; 2000WO-CN00502.

30-NOV-1999; 99CN-0124179.

(BIOR-) BIROAD GENE DEV LTD SHANGHAI.

Mao Y, Xie Y;

WPI; 2001-397944/42.

P-PSDB; AAB99224.

Isolated human brain-expressed X-linked polypeptide used to diagnose
and treat of dysembryoplasia, hereditary diseases, cancer, tumor,
deafness and X-chromosome-binding mental retardation

Claim 5; Page 22; 30pp; Chinese.

CC The present sequence is the coding sequence for a human brain-expressed
 CC X-linked protein (hhex). hhex and its coding sequence are useful in the
 CC diagnosis and treatment of dysembryoplasia, hereditary diseases, cancer,
 CC tumours, deafness, X-chromosome-binding mental retardation and
 CC lissencephalous disease. hhex is also useful for screening mimics,
 CC agonists, or inhibitors, and in peptide fingerprinting identification.
 CC hhex coding sequence can be used as primers or probes, or in producing
 CC gene chips or microarrays.

CC Sequence 792 BP; 214 A; 172 C; 219 G; 187 T; 0 other;

alignment_scores:
 Quality: 94.00 Length: 128
 Ratio: 1.362 Gaps: 2
 Percent Similarity: 53.906 Percent Identity: 28.906

alignment_block:
 US-09-327-750D-35 x AAH45143 ..

Align seg 1/1 to: AAH45143 from: 1 to: 792

```

14 GULYASPLYSLYSASPLYSARGLYLYS ALASERLYGLNSERG 30
   |||.....|.....|.....|.....|.....|.....|
187 GAACAGCACTAAACAGCTCAGCATGCAAAATGCCAACCAAGAAATG 236
   |||.....|.....|.....|.....|.....|.....|
30 LUGLUGLUPROHSHISLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 42
   |||.....|.....|.....|.....|.....|.....|
237 AAGAAAG.....GAGCAAGTCTCTATTAAGGAGGCCCTTG 274
   |||.....|.....|.....|.....|.....|.....|
43 .....LSPROGLYLYSVALNR 49
   |||.....|.....|.....|.....|.....|.....|
275 GCCCTCCTTGATGCTGCTGATCTGCTGCTGAGAGAAATGCTAG 324
   |||.....|.....|.....|.....|.....|.....|
49 GARG..LYSVALARGARGLEUALPROASNPHELEUTRPAALILEPRO 65
   |||.....|.....|.....|.....|.....|.....|
325 GCGGTCGCCGCTTAGCAGCCATCCGATAGATAGATAGATAGATAG 374
   |||.....|.....|.....|.....|.....|.....|
65 snARHISVALASP.....ARGASNGLUGLUGLUGLUGLUGLUGLUG 76
   |||.....|.....|.....|.....|.....|.....|
375 ATAGGCTTGAGAACACACAGCAGCAGATGAGAGAGAAATAGGAAG 424
   |||.....|.....|.....|.....|.....|.....|
77 VALGLYARGPHEVALGLINGLYTHGLUVALYLSARGLYSTHTRGL 93
   |||.....|.....|.....|.....|.....|.....|
425 ATGGGGAGGGGGGTGAGACAG.....CTGATGAGAAAGCTGAGGGA 465
   |||.....|.....|.....|.....|.....|.....|
93 UGNGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 107
   |||.....|.....|.....|.....|.....|.....|
466 AAGGAGTGTGAGTATAGTCTGCGGGCAGTCAGACCTGACCCCTCACC 515
   |||.....|.....|.....|.....|.....|.....|
108 ..ASPASNHISLYRASPHECYLSLEULEPRO 117
   |||.....|.....|.....|.....|.....|.....|
516 ATGACCATCATGATGATGATTGGCTTATAGCCC 547
   |||.....|.....|.....|.....|.....|.....|

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT: AAC01003

seq_documentation_block:

ID AAC01003 standard; cDNA; 401 BP.

AC AAC01003:

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 1001.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

OS Homo sapiens.

XX EPI033401-A2.

XX EPI033401-A2.

PD 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX WPI: 2000-500381/45.
 XX P-PSDB: AAG00997.
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1: SEQ ID 1001; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

CC Sequence 401 BP; 98 A; 98 C; 148 G; 57 T; 0 other;

alignment_scores:
 Quality: 92.00 Length: 52
 Ratio: 2.359 Gaps: 2
 Percent Similarity: 75.000 Percent Identity: 42.308

alignment_block:
 US-09-327-750D-35 x AAC01003 ..

Align seg 1/1 to: AAC01003 from: 1 to: 401

```

1 METALASERLYSLYSLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 16
   |||.....|.....|.....|.....|.....|.....|
238 ATGAGATCCAAAGAGAGAACTAGCGCAACATCTCAACGGGAAAAATGC 287
   |||.....|.....|.....|.....|.....|.....|
16 PLYSLYASPLYSARGLYLYS..LYSALASERLYSGLSERGGLUGLUG 32
   |||.....|.....|.....|.....|.....|.....|
288 CCACACAGCAAAACGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 337
   |||.....|.....|.....|.....|.....|.....|
32 LUPROHSHISLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUGLUG 48
   |||.....|.....|.....|.....|.....|.....|
338 AATCCCGCATTTGGGAGGGGTGAGGAGGAGGAGGAGGAGGAGGAG 387
   |||.....|.....|.....|.....|.....|.....|
49 ARGARG 50
   |||.....|.....|.....|.....|.....|.....|
388 AGCGCG 393
   |||.....|.....|.....|.....|.....|.....|

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT: AA120201

seq_documentation_block:

ID AA120201 standard; DNA; 276 BP.

AC AA120201:

DT 12-OCT-2001 (first entry)

XX Probe #10134 for gene expression analysis in human cervical cell sample.

Sequence	Strd Orig	EScore	Len	Documentation
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-931-999-4 +	80.00	137.69	5.09	6755 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-338-997-179 -	74.50	103.29	420.15	56520 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-996-306-1 -	71.50	96.55	997.04	56516 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-338-997-1 -	71.50	96.55	997.04	56516 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-961-083-37 +	71.00	134.11	8.06	1360 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-09-258-373-21 +	71.00	130.78	12.36	1875 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-841-483-3 +	71.00	124.33	28.27	3450 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-382-911-3 +	71.00	124.33	28.27	3450 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-841-483-3 +	71.00	124.33	34.97	4094 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-382-911-5 +	71.00	122.67	34.97	4094 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-08-402-282-3 +	71.00	108.75	208.65	15664 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-08-508-004-3 +	71.00	108.75	208.65	15664 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-08-402-066-3 +	71.00	108.75	208.65	15664 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-08-403-068-3 +	71.00	108.75	208.65	15664 !
/cgn2_6.ptodata/2/1na/PTGUS.COMB.seq:PC8-US93-03077-4 +	70.50	133.93	8.25	1242 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-08-198-446B-9 +	70.50	131.00	12.02	1647 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-870-693-9 +	70.50	131.00	12.02	1647 !
/cgn2_6.ptodata/2/1na/PTGUS.COMB.seq:PC8-US93-03077-2 +	70.50	123.85	30.05	3279 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-103-537-34 +	70.50	120.14	48.38	4689 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-103-537-5 +	70.50	98.76	750.67	36778 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-09-320-878-19 +	70.50	98.29	797.98	38506 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-103-840A-2 -	70.50	47.97	3.6e+05	4403765 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-403-852D-2 +	69.50	131.47	11.32	1268 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-510-646B-2 +	69.50	131.47	11.32	1268 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-231-818-2 +	69.50	131.47	11.32	1268 !
/cgn2_6.ptodata/2/1na/5B.COMB.seq:US-08-403-852D-1 +	69.50	116.45	77.72	5392 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-510-646B-1 +	69.50	116.45	77.72	5392 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-231-818-1 +	69.50	116.45	77.72	5392 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-103-840A-2 -	69.50	46.85	4.0e+05	4403765 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-103-840A-1 -	69.50	46.83	4.0e+05	4411529 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-948-705-3 +	69.00	130.10	13.48	1298 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-09-120-365-89 +	69.00	127.07	19.89	1738 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-515-039-89 +	69.00	127.07	19.89	1738 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-09-120-365-4 +	69.00	126.49	21.44	1839 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-09-515-039-4 +	69.00	126.49	21.44	1839 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-826-246-1 +	69.00	125.86	23.23	1953 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-08-944-435-1 +	69.00	125.86	23.23	1953 !
/cgn2_6.ptodata/2/1na/6A.COMB.seq:US-09-126-640-1 +	69.00	125.86	23.23	1953 !
/cgn2_6.ptodata/2/1na/6B.COMB.seq:US-08-925-588-1 +	69.00	125.86	23.23	1953 !
/cgn2_6.ptodata/2/1na/5A.COMB.seq:US-07-807-043B-5 -	68.50	115.63	80.30	4698 !


```
; LOCATION: 13617..13636
; OTHER INFORMATION: downstream amplification primer 4-72, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 13547..13564
; OTHER INFORMATION: upstream amplification primer 4-73
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 13962..13981
; OTHER INFORMATION: downstream amplification primer 4-73, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 15994..16011
; OTHER INFORMATION: downstream amplification primer 99-610
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 16463..16480
; OTHER INFORMATION: upstream amplification primer 99-610, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 17304..17324
; OTHER INFORMATION: downstream amplification primer 99-609
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 17814..17832
; OTHER INFORMATION: upstream amplification primer 99-609, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 18008..18027
; OTHER INFORMATION: upstream amplification primer 4-90
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 18423..18442
; OTHER INFORMATION: downstream amplification primer 4-90, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 18699..18716
; OTHER INFORMATION: downstream amplification primer 99-607
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 19164..19182
; OTHER INFORMATION: upstream amplification primer 99-607, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 22589..22609
; OTHER INFORMATION: downstream amplification primer 99-602
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 23111..23129
; OTHER INFORMATION: upstream amplification primer 99-602, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 25098..25118
; OTHER INFORMATION: downstream amplification primer 99-600
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 25657..25674
; OTHER INFORMATION: upstream amplification primer 99-600, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 26537..26557
; OTHER INFORMATION: downstream amplification primer 99-598
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 27022..27040
; OTHER INFORMATION: upstream amplification primer 99-598, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 32262..32281
; OTHER INFORMATION: downstream amplification primer 99-592
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 32823..32841
```

```
; OTHER INFORMATION: upstream amplification primer 99-592, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 34215..34233
; OTHER INFORMATION: upstream amplification primer 99-217
; FEATURE:
alignment_scores:
    Quality: 74.50      Length: 82
    Ratio: 1.656       Gaps: 2
    Percent Similarity: 54.878    Percent Identity: 30.488
alignment_block:
US-09-327-750D-35 x US-09-338-907-179/rev ..
Align seg 1/1 to reverse of: US-09-338-907-179 from: 1 to: 56520
14 GluLysAspLysLysAspLysArgGlyGlyLysAlaSerLysGlnSerG1 30
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39085 GAAGCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 39036
30 uclugluProHisLeuGluGluValGluAsnLysLysProGlyGlyA 47
: : : : : : : : : : : : : : : : : : : : : : : :
39035 AACGAGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38986
47 snValArgGlyValArgGlyLeuValProAsnPheLeuTrpAlaIle 63
: : ||| ||| |||||
38985 GAGAGAGAGCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38963
64 ProAsnArgHisValAspArgAsnGluGlyGlu.....As 76
: : : : : : : : : : : : : : : : : : : : : : : :
38962 .....AGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 38931
76 pValcArgPheValValGlnGlyThrGluValLysArgLysThr 91
: : : : : : : : : : : : : : : : : : : : : : : :
38930 AATGGGCATACTGAGGCTGAAGGAGCAACAACTCTAGAAGTAGT 38885
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-996-306-1
seq_documentation_block:
; Sequence 1, Application US/08996306
; Patent No. 5945522
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Chumakov, Ilya
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Prostate cancer gene
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 901 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,306
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: GENSET.018A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 1:
```



```
79 gpheValValGlnGlyThrGluValValLysArgLysThr 91
      ::::::::::::::::::::
39922 TACTGAGCTGAGGACAAACATCTAGAGTAGT 38886
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-961-083-37
seq_documentation_block:
; Sequence 37, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-37

alignment_scores:
      Quality: 71.00      Length: 104
      Ratio: 1.224      Gaps: 4
      Percent Similarity: 55.769      Percent Identity: 25.962

alignment_block:
US-09-327-750D-35 x US-08-961-083-37 ..
Align seg 1/1 to: US-08-961-083-37 from: 1 to: 1360
5 LysGlnValLeuAspLeuThrValGluLysAspLysLysArg 21
      ::::::::::::::::::::
413 AAAAAGCGGAGCTTGAACTA...GTAAGAGGAGGAGCTAGGAACTCG 459
21 gGlyGlyLysAlaSerLysGlnSerGluGluProHisLeuGlu 38
      ::::::::::::::::::::
460 AAACAGGAGAAAGTTAGCAAGCAAA.....GCGG 491
38 luValGluAsnLysProGlyGlyAsnValArgLysValArg 54
      ::::::::::::::::::::
492 AAGTTAGAGTAAAGAACTGAGGCTACAGGTTAGAAAATCAAG... 538
55 LeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspArgAs 71
      ::::::::::::::::::::

539 .....ACAGATCGTAA 549
71 nGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrGluValL 88
      ::::::::::::::::::::
550 AAAAGCAGAGAAGAAGCTAAACGAAAGACGACGAGAGAGATAAGTTA 599
88 ysArgLysThrThrGluGln....GlnValArgProTyrArgArgPheArg 103
      ::::::::::::::::::::
600 AAGAAAAACCAAGCTGAACCAACCAACGCGCGCTCCCAAAAGCAGAA 649
104 ThrProGluPro 107
      ::::::::::::::::::::
650 AAACAGAGCTCCA 661
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-258-373-21
seq_documentation_block:
; Sequence 21, Application US/09258373
; Patent No. 6150110
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Xiao, Sheng
; TITLE OF INVENTION: HMG(IY)-LAMA4* FUSION ONCOGENE/
; FILE REFERENCE: B0801/7135/ERP
; CURRENT APPLICATION NUMBER: US/09/258,373
; EARLIER FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 60/076,401
; EARLIER FILING DATE: 1998-02-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 1875
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-09-258-373-21

alignment_scores:
      Quality: 71.00      Length: 118
      Ratio: 1.224      Gaps: 7
      Percent Similarity: 49.153      Percent Identity: 29.661

alignment_block:
US-09-327-750D-35 x US-09-258-373-21 ..
Align seg 1/1 to: US-09-258-373-21 from: 1 to: 1875
2 AlaSerLysLysGlnValLeuAspLeuThrValGluLysAspLysLy 18
      ::::::::::::::::::::
244 AGCTCGAAGTCCAGCCAGCCCTTGGCTCCAAAGCAGAGAAAGGACGAC 293
18 sAspLysArgGly...GlyLysAlaSerLysGlnSerGluGluProH 34
      ::::::::::::::::::::
294 TGAGAACGGGGCGGGGCGGAGCGCCGCGCAAGCAGCGCTCCGAGAGAGCCA 343
34 ishIsLeuGluGluValGluAsnLysLysPro.GlyGlyAsnValArg 50
      ::::::::::::::::::::
344 GC.....GAAGTCCCAACACCTAAGAGACCTCGGGCGCGCAAGGGA 387
50 gLysValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArgH 67
      ::::::::::::::::::::
388 AGCAAAACCAAGGTGCTGCCAAGA.....CCCGAAAC 422
67 isValAspArgAsnGluGlyGlyGluAspValGlyArgPheValGln 83
      ::::::::::::::::::::
423 CACCACAACCTCCAGGAAGGAAAC.....CAA 448
84 GlyThrGluValLys.....ArgLysThrThrGluGlnGlnVa 96
      ::::::::::::::::::::
449 GGGCAGACCCCAAAACTGGAGAGAGAGAGAGGAGGCGCATCTCGCAG 498
96 lArgProTyrArgArgPheArgThrPro.....GluProAspAsnHist 111
```


alignment_scores:	Quality:	71.00	Length:	82
-------------------	----------	-------	---------	----

```

; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"

```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; US-08-402-282-3
;
alignment_scores:
; Quality: 71.00 Length: 71
; Ratio: 1.775 Gaps: 4
; Percent Similarity: 56.338 Percent Identity: 33.803
;
alignment_block:
; US-09-327-750D-35 x US-08-402-282-3
;
Align seg 1/1 to: US-08-402-282-3 from: 1 to: 15664
;
49 AtgAtgLySValArgArgLeuValProAsnPhelutrpAlaileProas 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAATCGGGTGGCGCGCAAGCGCGCGCGCGCGCGCGCGCGG.....CCGAC 13479
;
65 nArg.His.....ValAspArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:|||||:
13480 GCGCTCACCCCGAACCGCGACCTCGAGGTGCGATCGGTTCCGGGGA 13529
;
74 GlyGluAspValGlyArgPheValGlnGlyThrGluValLysArgly 90
|||||:|||||:|||||:|||||:|||||:|||||:
13530 GGAATGAT.....CGAAGCCTACCGCGCGCGCGCGCGCGCGCGG 13573
;
90 sThrThrGluGlnGlnValArgProTyrArgArgPheArgThrProGluP 107
|||||:|||||:|||||:|||||:|||||:|||||:
13574 GCGAGCAGCAGGAATCGACGCCCGCCAGCAGCAGCATCGTGACGTCGATC 13623

```

```
107 roaspaashis 110
|| :|||
13624 CTGTCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-508-004-3

seq documentation block:
; Sequence 3, Application US/08508004
; Patent No. 5582969
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Bayer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/508,004
; FILING DATE: 27-JUL-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/402,282
; FILING DATE: 10-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"

;
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading
; OTHER INFORMATION: frame"
```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 11917..12741
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12748..14499
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14771..15154
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15154..15426
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15429..15664
OTHER INFORMATION: /function= "potential open reading"
OTHER INFORMATION: frame"
US-08-508-004-3
alignment_scores:
Quality: 71.00 Length: 71
Ratio: 1.775 Gaps: 4
Percent Similarity: 56.338 Percent Identity: 33.803
alignment_block:
US-09-327-750D-35 x US-08-508-004-3
Align seg 1/1 to: US-08-508-004-3 from: 1 to: 15664
49 ArgArgLysValArgArgLeuValProAspPheLeuTIPAlaIleProAs 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13436 CGAATCGGGTGGGGCAGGCGCCCGCGATGCCGTGG.....CCGAC 13479
65 nArg.His.....ValaspArgAsnGluGly 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13480 GCGCTACCCGCGAAGCCGCGACCTCGAGGTCCGTCGATCGGTCCGGGA 13529
74 GlyGluAspValGlyArgPheValValGlnGlyThrGluValLysArgly 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13530 GGGATGAT.....CGAAGCCTACCGCGGCGACACGACGACAAAGCGCG 13573
90 sThrThrGluGlnGlnValArgProTyrArgPheArgPheArgThrProGlu 107
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13574 GCGACGACGGGATCGAGCGGCGCCGACGACGACATCGTCGATCGATC 13623
107 roAspAsnHis 110
||:|||||
13624 CTGTCAGCAT 13634
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-066-3
seq_documentation_block:
; Sequence 3, Application US/08402066
; Patent No. 5612182
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Bayer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS COMPLEX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/402,066
; FILING DATE:
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugitt, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame"
; FEATURE:

```

```
;
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding"
; OTHER INFORMATION: sequence
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6405..7770
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10244
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
```

```
;
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; US-08-402-066-3

alignment_scores:
  Quality: 71.00      Length: 71
  Ratio: 1.775      Gaps: 4
  Percent Similarity: 56.338      Percent Identity: 33.803

alignment_block:
  US-09-327-750D-35 x US-08-402-066-3
  Align seg 1/1 to: US-08-402-066-3 from: 1 to: 15664
    49 ArgArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAs 65
    |||:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    13436 CGAATCGGTGGCGGCAAGGCGCGCGATGCGCGTGG.....CCGAC 13479
    65 nArg.His.....ValaspArgAsnGluGly 73
    |||:|||:|||||:|||||:|||||:|||||:|||||:|||||:
    13480 GCGCTACCCCGAACCGCGACCTCGAGGTCGATCGGTTCGGGGA 13529
    74 GlycluAspValGlyArgPheValValGlnGlyThrGluValLysArgLy 90
    |||:|||:|||||:|||||:|||||:|||||:|||||:|||||:
    13530 GGGATGAT.....CGAGCCTACCGCGGCGGACAGAGAAAAACGGCG 13573
    90 sThrThrGluGlnValArgProTyrArgArgPheArgThrProGluP 107
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:
    13574 GCGACGACGACGGAATCGACGCGGCCACAGCAGCATCGTACGTCGATC 13623
    107 roAspAsnHis 110
    |||:|||:|||||
    13624 CTGTTTCAGCAT 13634

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-402-068-3

seq_documentation_block:
; Sequence 3, Application US/08402068
; Patent No. 5633159
; GENERAL INFORMATION:
; APPLICANT: Pearson, Robert E.
; APPLICANT: Dickson, Julie A.
; APPLICANT: Hamilton, Paul T.
; APPLICANT: Little, Michael C.
; APPLICANT: Beyer Jr., Wayne F.
; TITLE OF INVENTION: MYCOBACTERIOPHAGE SPECIFIC FOR THE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/402,068
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3283
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15664 base pairs
```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 222..425
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 451..747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 747..1109
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1109..2014
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2034..2747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2747..3109
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3109..3444
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3444..3728
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 3731..4855
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4855..5376
; OTHER INFORMATION: /function= "potential coding"
; OTHER INFORMATION: sequence
; OTHER INFORMATION: /product= "L5 gp37 homolog"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5382..5747
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5837..6307
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 6403..7770
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7770..8006
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; OTHER INFORMATION: /function= "potential open reading"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8033..8236
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8244..9443
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9450..10274
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 10371..10586
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11115..11786
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 11917..12741
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 12748..14499
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 14771..15154
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15154..15426
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 15429..15664
; OTHER INFORMATION: /function= "potential open reading"
; OTHER INFORMATION: frame
; US-08-402-068
; alignment_scores:
;   Quality: 71.00 Length: 71
;   Ratio: 1.775 Gaps: 4
;   Percent Similarity: 56.338 Percent Identity: 33.803
; alignment_block:
;   US-09-327-750D-35 x US-08-402-068-3 ..
;   Align seg 1/1 to: US-08-402-068-3 from: 1 to: 15664
;   49 ArgArgLysValArgArgLeuValProAsnPhleLeuTrpAlaIleProAs 65
;   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
;   13436 CGAATCGGTGGCGGCAAGGCGCGCGATGCGCGTGG.....CCGAC 13479
;   65 nArg.His.....ValAspArgAsnGluGly 73
;   :|||:|||||:|||||:|||||:|||||:|||||:|||||:
;   13480 GCGCTCACCCCGAACCGCGACCTCGAGGTCCGTCGATCGGTCCGGGA 13529

```


OM of: US-09-327-750D-35 to: GenEmbl:* out_format : pfs
 Date: Mar 11, 2002 3:35 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEP=frame+ p2n.model -DEV=xlh
 -O=/cgn2.1/USPTO.spool/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
 -DB=GenEmbl -QFT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPL=0.000 -XGAPEXT=0.000 -XGAPOP=4.500
 -GAPEXT=0.050 -XGAPOP=10.000 -YGAPEXT=0.500 -YGAPOP=6.000
 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -LIST=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi
 -ALIGN=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -NLLEN=0 -MODE=LOCAL -OUTFT=pfs -NORM=ext -HEAPSIZ=500
 -NCPU=6 -ICPU=3 -LONGLOG -DEV=US09327750 -CGNL_1.8673
 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-327-750D-35
 Query length: 117
 Database: GenEmbl:*
 Database sequences: 1472140
 Database length: 341344837
 Search time (sec): 4557.230000

score_list:

Sequence	Strd Orig	zScore	Escore	Len	Documentation
gb_pr:AK000959	+ 355.50	626.28	1.3e-26	1229	AK000959 Homo sapiens CDNA FLJ
gb_pat:AX100231	+ 355.50	625.44	1.5e-26	1364	AX100231 Sequence 32 from Pat
gb_pr:HS635G19	+ 355.50	593.78	8.5e-25	69648	AL035494 Human DNA sequence
gb_pr:BC003190	+ 175.00	306.91	9.1e-09	793	BC003190 Homo sapiens, p75NTR-3
gb_pr:HM00GC	+ 175.00	305.97	9.1e-09	891	M38188 Human unknown protein fr
gb_pr:AF187064	+ 175.00	305.97	9.1e-09	891	M38188 Human unknown protein fr
gb_to:AF187066	+ 154.00	270.34	8.8e-07	700	AF187066 Mus musculus p75NTR-as
gb_to:G72708	+ 150.00	267.28	1.1e-06	854	AF097440 Mus musculus brain exp
gb_to:AF187065	+ 149.50	264.70	1.3e-06	421	G72708 MARC 4953-4954:195139031
gb_to:HSV351F8	+ 138.00	260.08	0.0026	519	AF187065 Rattus norvegicus p75N
gb_sts:G35294	+ 120.50	213.50	0.0013	477	G35294 Human STS SHGC-37409, se
gb_pr:HS714B7	+ 107.50	185.74	0.0454	835	AF097438 Mus musculus brain exp
gb_to:AF051347	+ 107.50	185.71	0.0456	835	AF051347 Mus musculus REX-3 mRN
gb_to:AF097439	+ 102.00	176.39	0.1275	2289	AF097437 Mus musculus Bex1 pro
gb_to:AF231053	+ 99.50	172.35	0.2528	785	AF097439 Mus musculus brain exp
gb_pat:AX078272	+ 99.50	170.84	0.3070	744	AF251053 Homo sapiens X-linked
gb_pr:HSV870H8	+ 99.50	142.25	12.02	898	AX078272 Sequence 76 from Pat
gb_pr:AL133348	+ 99.50	140.16	15.70	31321	AL133348 Human DNA sequence fr
gb_htg:HS080B1	+ 99.50	140.07	15.88	40584	AL133348 Human DNA sequence fr
gb_pr:HS753P9	+ 98.50	129.88	58.71	116561	AL022169 Homo sapiens chtrmos
gb_pr:AF195953	+ 98.50	125.19	107.09	208618	AF195953 Homo sapiens membra
gb_htg:AC012947	+ 95.00	121.94	162.57	143599	AC012947 Drosophila melanoga
gb_in:AC009183	+ 95.00	119.60	219.49	193055	AC009183 Drosophila melanoga
gb_in:AC010124	+ 95.00	119.54	221.08	193398	AC010124 Drosophila melanoga
gb_pr:AF183416	+ 94.00	118.32	258.52	225038	AF183416 Homo sapiens ovarian g
gb_pr:AF237783	+ 94.00	162.02	0.9513	791	AF237783 Homo sapiens uncharac
gb_pr:HS198P4	+ 94.00	161.65	0.9972	828	AF237783 Homo sapiens uncharac
gb_pr:AC002478	+ 94.00	131.35	48.61	35714	AC002478 Human BAC clone GS1
gb_htg:AC012232	+ 91.00	120.55	194.16	70113	AC002478 Human BAC clone GS1
gb_ba:MXAPRZ	+ 91.00	113.22	497.28	174337	AC012232 Homo sapiens clone
gb_pr:AP000265	+ 90.50	146.47	6.99	2507	J04157 M.xanthus fra, fraB an
gb_pr:AP000034	+ 90.50	124.52	116.65	38304	AP000265 Homo sapiens genom
gb_pr:AP000101	+ 90.50	116.80	314.19	100000	AP000034 Homo sapiens genom
gb_pr:AP000177	+ 90.50	116.80	314.19	100000	AP000101 Homo sapiens genom
gb_htg:AC026930	+ 90.50	114.66	413.39	130443	AC026930 Homo sapiens chrom

gb_pr:AC005295 + 90.50 114.62 415.44 131070 ! AC005295 Homo sapiens Xp2
 gb_pr:AP001713 + 90.50 106.95 1.1e+03 340000 ! AP001713 Homo sapiens gen
 gb_ba:AE008371 - 90.00 133.00 39.36 11967 ! AE008371 Agrobacterium tum
 gb_pr:U72787 + 90.00 129.48 61.75 18510 ! U72787 Homo sapiens cosmid
 gb_pr:AL353806 + 90.00 115.06 392.66 111051 ! AL353806 Human DNA sequen

seq_name: gb_pr:AK000959

seq_documentation_block:

LOCUS AK000959 1229 bp mRNA PRI 22-FEB-2000
 DEFINITION Homo sapiens CDNA FLJ10097 f1s, clone HEMBA1002458, weakly similar
 to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.
 ACCESSION AK000959
 VERSION AK000959.1 GI:7021945
 KEYWORDS oligo clapping; f1s (full insert sequence).
 SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
 mRNA, clone_lib:HEMBA1002458.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (sites)
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
 Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
 Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
 Wakamatsu,A., Nakamura,Y., Negahari,K., Masuno,Y. and Oshima,A.
 NEDO human cDNA sequencing project
 Unpublished (2000)

JOURNAL

2 (bases 1 to 1229)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao

AUTHORS

Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (16-FEB-2000) to the DDBJ/EMBL/GenBank databases. Takao

JOURNAL

Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
 Tel:81-438-52-3951, Fax:81-438-52-3952)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing; Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.

FEATURES

Location/Qualifiers
 1..1229

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HEMBA1002458"

/dev_stage="embryo, 10 weeks"

/tissue_type="whole embryo, mainly head"

/note="cloning vector: pME18SFL3"

191..553

/note="unnamed protein product"

/codon_start=1

/protein_id="BA091443.1"

/db_xref="GI:7021945"

/translation="MESKELANNLNGENAGQENGGQAPTONEEESRLHGGEGQ
 APGGNIRGRVRRVLPNFRWAIPIRHHNEARDVDFRVGOMIEIKRTRQOMRHY
 MRFTQPDNHYDFCLIP"

BASE COUNT 351 a 243 c 296 g 339 t

ORIGIN

alignment_scores:

Quality: 355.50 Length: 120

Ratio: 3.665 Gaps: 3

Percent Similarity: 80.833 Percent Identity: 59.167

alignment_block:

US-09-327-750D-35 x AK000959

Align seg 1/1 to: AK000959 from: 1 to: 1229

details see [http://bacpac.med.buffalo.edu/ VECTOR: pCYPAC2](http://bacpac.med.buffalo.edu/VECTOR: pCYPAC2)
 IMPORTANT: This sequence is not the entire insert of clone 635G19.
 It may be shorter because we only sequence overlapping sections
 once, or longer because we arrange for a small overlap between
 neighbouring submissions.
 The true left end of clone U65A4 (281014) is at 69549 in this
 sequence. The true right end of clone U101D3 (285937) is at 100 in
 this sequence.

FEATURES

source	Location/Qualifiers
	1. 69648
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="X"
	/map="q22.1-22.3"
	/clone="RP4-635G19"
	/clone_id="RPI-4"
	439. 483
repeat_region	/note="TAR1 repeat: matches 707. 749 of consensus"
repeat_region	501. 661
repeat_region	/note="AluJb repeat: matches 137. 311 of consensus"
	662. 897
repeat_region	/note="118 copies 2 mer aa 70% conserved"
	915. 1331
repeat_region	/note="MSA repeat: matches 1. 419 of consensus"
	1334. 1517
repeat_region	/note="92 copies 2 mer ga 77% conserved"
	1927. 2020
repeat_region	/note="MIR repeat: matches 109. 217 of consensus"
	2066. 2262
repeat_region	/note="L1ME2 repeat: matches 5825. 6022 of consensus"
	2263. 2574
repeat_region	/note="AluSp repeat: matches 1. 313 of consensus"
	2575. 2819
repeat_region	/note="L1ME2 repeat: matches 5583. 5825 of consensus"
	2821. 3039
repeat_region	/note="Alu repeat: matches 1. 311 of consensus"
	3104. 3384
repeat_region	/note="L1ME2 repeat: matches 5304. 5591 of consensus"
	3391. 4007
repeat_region	/note="L1MB1 repeat: matches 5558. 6162 of consensus"
	4012. 4137
repeat_region	/note="Trig2a repeat: matches 11. 123 of consensus"
	4141. 4252
repeat_region	/note="L1 repeat: matches 2920. 3031 of consensus"
	4265. 4350
repeat_region	/note="HY4 repeat: matches 1. 87 of consensus"
	4350. 4678
repeat_region	/note="MER74A repeat: matches 28. 369 of consensus"
	4914. 5095
repeat_region	/note="MER5A repeat: matches 8. 189 of consensus"
	5224. 5555
repeat_region	/note="L2 repeat: matches 2423. 2746 of consensus"
	6040. 6086
repeat_region	/note="L2 repeat: matches 1676. 1720 of consensus"
	7413. 7478
repeat_region	/note="Alu repeat: matches 239. 304 of consensus"
	7628. 7928
repeat_region	/note="AluYb8 repeat: matches 1. 316 of consensus"
	7952. 7987
repeat_region	/note="18 copies 2 mer aa 81% conserved"
	7990. 8305
repeat_region	/note="AluJb repeat: matches 1. 309 of consensus"
	8436. 8538
repeat_region	/note="MIR repeat: matches 99. 212 of consensus"
	10836. 11442
repeat_region	/note="L1MD repeat: matches 1357. 1972 of consensus"
	11443. 11571
repeat_region	/note="FLAM_A repeat: matches 2. 129 of consensus"
	11572. 11946
repeat_region	/note="L1MD repeat: matches 977. 1357 of consensus"
	11947. 12245
repeat_region	/note="AluY repeat: matches 1. 301 of consensus"
	12246. 13098
	/note="L1MD repeat: matches 275. 977 of consensus"
	13094. 13223
	/note="L1MD repeat: matches -17. 113 of consensus"
	13523. 13820
	/note="L1MD repeat: matches 177. 505 of consensus"
	13823. 13932
	/note="55 copies 2 mer ct 77% conserved"
	13936. 13977
	/note="L1MD repeat: matches 201. 241 of consensus"
	13978. 14240
	/note="AluSg1 repeat: matches 1. 270 of consensus"
	14245. 14534
	/note="AluSc repeat: matches 1. 290 of consensus"
	14535. 14721
	/note="L1MD repeat: matches 2. 201 of consensus"
	14724. 15604
	/note="L1MC/D repeat: matches 4646. 5565 of consensus"
	15605. 15909
	/note="AluX repeat: matches 24. 312 of consensus"
	16069. 16363
	/note="AluX repeat: matches 1. 291 of consensus"
	16482. 16641
	/note="AluJb repeat: matches 136. 295 of consensus"
	16671. 16953
	/note="AluX repeat: matches 1. 284 of consensus"
	17164. 17280
	/note="L1P repeat: matches 4499. 4619 of consensus"
	17326. 17377
	/note="26 copies 2 mer at 79% conserved"
	17704. 17863
	/note="AluSc repeat: matches 118. 295 of consensus"
	17864. 18175
	/note="AluXb8 repeat: matches 1. 318 of consensus"
	18176. 18298
	/note="AluSc repeat: matches 1. 118 of consensus"
	18848. 18873
	/note="13 copies 2 mer tg 100% conserved"
	18875. 19171
	/note="AluSg repeat: matches 1. 295 of consensus"
	19183. 19232
	/note="25 copies 2 mer ac 100% conserved"
	19246. 19533
	/note="AluY repeat: matches 5. 292 of consensus"
	19539. 19837
	/note="MER74B repeat: matches 61. 368 of consensus"
	20016. 20194
	/note="MER74A repeat: matches 295. 487 of consensus"
	20394. 20692
	/note="AluSg repeat: matches 1. 302 of consensus"
	20701. 20809
	/note="L1R1 repeat: matches 682. 785 of consensus"
	20877. 21061
	/note="L1R1 repeat: matches 1. 195 of consensus"
	21186. 21636
	/note="L1R29 repeat: matches 3. 498 of consensus"
	21605. 21737
	/note="MER34 repeat: matches 409. 543 of consensus"
	21771. 22076
	/note="AluY repeat: matches 2. 308 of consensus"
	22878. 22974
	/note="MIR repeat: matches 92. 191 of consensus"
	23452. 23582
	/note="FLAM_C repeat: matches 1. 131 of consensus"
	23583. 23624
	/note="21 copies 2 mer ta 76% conserved"
	23735. 24041
	/note="AluSg1 repeat: matches 1. 308 of consensus"
	24045. 24072
	/note="14 copies 2 mer tt 89% conserved"
	25274. 25582
	/note="AluJb repeat: matches 1. 289 of consensus"
	25583. 25747
	/note="match: GSS B63622"
misc_feature	

/note="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75129.1"
/db_xref="GI:8452894"
/translation="NANHOENEMEOPMONGEDRPLGGEGHQPAGNRGQARRLA
PNPRAIPNRQINDCMGGDDMEIFMEEMREIRKRLQRLNCLRLMGELSNHHD
HDEFCIMP"
BASE COUNT 251 a 182 c 224 g 234 t
ORIGIN
alignment_scores:
Quality: 175.00 Length: 121
Ratio: 2.273 Gaps: 4
Percent Similarity: 63.636 Percent Identity: 36.364
alignment_block:
US-09-327-750D-35 x AF187064 ..
Align seg 1/1 to: AF187064 from: 1 to: 891
4 LysLysGlnValIleLeu...AspLeuThrValGluLysAspLysLysAs 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 36
297 AAAAAAATCTCATCTGCAATATTCACAGGAAACGAGAGATGGA 346
19 pLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProHisL 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 36
347 GCAG.....CCTATGCAGATGGAGGAGGAGCCGCCCTT 381
36 eGluGluValGluLysLysProGlyGlyAsnValArgArgLysVal 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 52
382 TGGAGAGGTGAGGACGACGAGTGCAGGAAATCGACGGGACAGGCT 431
53 ArgArgLeuValProAspPheLeuTrpAlaIleProAsnArgHisVal 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 69
432 CGCCGACTTGGCCCTAAATTTGGATGGCCCATACCAATGAGGAGATCA 481
69 pArgAsnGluGly.....GlyGluAspValGlyArgPheValValGln 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 84
482 TGATGGATGGTGGGAGATGAGATGATGATAATATTCATGAGGAGA 531
84 lYThrGluValLysArgLysThrThrGluGlnGlnValArgProTyArg 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 100
532 TGAGAGAAATCAGAGAAACTAGGAGCTGCAGTTGAGGAATTGCTG 581
101 ArgPheArgThrProGluProAspAsnHisTyrAsp.....Ph 113
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 113
582 CGTATCCTTATGGGAGCTCTCTAATCACCATGACCATCATGATGAAT 631
113 eCysLeuIlePro 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 117
632 TTGCTTATGCTT 644
seq_name: gb_ro:AF187066
seq_documentation_block:
LOCUS AF187066 700 bp mRNA ROD 11-JUN-2000
DEFINITION Mus musculus p75NTR-associated cell death executor (Nade) mRNA,
complete cds.
ACCESSION AF187066
VERSION AF187066.1 GI:8452897
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
AUTHORS Mukai,J., Hachiya,T., Shoji,Hoshino,S., Kimura,M.T., Nadano,D.,
Suvanto,P., Hanaoka,T., Li,Y., Irie,S., Greene,L.A. and Sato,T.A.
TITLE Nade, a p75NTR-associated cell death executor, is involved in
signal transduction mediated by the common neurotrophin receptor

p75NTR
J. Biol. Chem. 275 (23), 17566-17570 (2000)
20298829
2 (bases 1 to 700)
REFERENCE Mukai,J., Hachiya,T., Hoshino,S., Kimura,M., Nadano,D., Suvanto,P.,
Hanaoka,T., Li,Y., Irie,S. and Sato,T.
AUTHORS Direct Submission
TITLE Otolaryngology/Pathology, Columbia
JOURNAL Submitted (17-SEP-1999)
UNIVERSITY, 630 West 168th St., P&S 11-451, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..700
Source
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
1..700
gene
177..551
CDS
/gene="Nade"
/note="Involved in the common neurotrophin receptor
p75NTR-mediated signal transduction; Nade"
/codon_start=1
/product="p75NTR-associated cell death executor"
/protein_id="AAF75131.1"
/db_xref="GI:8452898"
/translation="MANVHOENEMEOPNQEQDRPVGGEGHQPAAANNHNNHNN
HNHRRGQARRLAIPNFRWAIPINRQNDLGGDDMEIFMEEMREIRKRLQRLNCL
LRLMGELSNHHDHDEFCIMP"
BASE COUNT 178 a 187 c 203 g 132 t
ORIGIN
alignment_scores:
Quality: 154.00 Length: 136
Ratio: 1.925 Gaps: 4
Percent Similarity: 58.824 Percent Identity: 30.147
alignment_block:
US-09-327-750D-35 x AF187066 ..
Align seg 1/1 to: AF187066 from: 1 to: 700
1 MetaSerLysLysGlnValIleLeuAspLeuThrValGluLysAspLys 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 17
150 ATAGCAACCAAGAAACAAATCTCATGCGCAATGTCACCAAGGAAAA 199
17 sLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluGluProH 34
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 34
200 CGAAGAGATG.....GAGCAGCCCTGCAGAAATGGACAGGAAGACC 240
34 iShLysLeuGluValGluAsnLysLysProGlyGlyAsn..... 47
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 47
241 GCCTGTGGGAGGAGGTGAGGGCCACCGCTGCTGCAAAACAAACAAAC 290
48ValArgArgLysValArgArg 54
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 54
291 AACAAACCAACCAATACCAACCAACCAACCAACCAACCAACCAACCA 340
54 gLeuValProAspPheLeuTrpAlaIleProAsnArgHisValAspArg 71
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 71
341 ACTTGGCCCTAACTCCGATGGGCCATCCCAACAGGAGGATGATGACG 390
71 snGluGly.....GlyGluAspValGlyArgPheValValGlnGlyThr 85
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 85
391 GGTGGGTGGAGATGGAGATGATATGAAATGTTTCATGGAGGAGATGAGA 440
86 GluValLysArgLysThrThrGluGlnGlnValArgProTyArgArgph 102
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 102
441 GAGATCCGGAGAAAGCTTGGAGGAGCTACAGCTGAGAAATGTTCTACGAT 490
102 eArgThrProGluProAspAsnHisTyrAsp.....PheCysL 115
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||| 115
491 CCTTATGGGAGCTGCTGAACCAACCAACCAACCAACCAACCAACCAAT 540

115 euilePro 117
 ||:|||||
 541 TTATGCT 548

seq_name: gb_ro:AF097440

seq_documentation_block:

LOCUS AF097440 854 bp mRNA ROD 13-APR-1999
 DEFINITION Mus musculus brain expressed x-linked protein 3 (Bex3) mRNA,
 complete cds.

ACCESSION AF097440
 VERSION AF097440.1 GI:4580593

KEYWORDS

SOURCE

ORGANISM

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

Erratum: [[published erratum appears in Hum Mol Genet 1999
 May;8(5):943]]

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .854
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /db_xref="dbEST:AA272375"
 /chromosome="X"
 /map="near Plp"
 /tissue_type="pooled organs"
 1. .854
 /gene="Bex3"
 172. .546
 /gene="Bex3"
 /codon_start=1
 /product="brain expressed x-linked protein 3"
 /protein_id="AA024431.1"
 /db_xref="GI:4580594"
 /translation="MANVHQENEMEOPLQNGQEDRPVGGEGHQPAANNNNHHNHN
 HNHRRGQARRLAPNFRWAIPIRQMDGLGGDDMMFMEIRKRLRELQRLNC
 LRILMGLSNHHDHDFCLMP"

BASE COUNT 237 a 212 c 228 g 177 t

ORIGIN

alignment_scores:

Quality: 154.00 Length: 136

Ratio: 1.925 Gaps: 4

Percent Similarity: 58.824 Percent Identity: 30.147

alignment_block:

US-09-327-750D-35 x AF097440 ..

Align seg 1/1 to: AF097440 from: 1 to: 854

1 MetAlaSerLysGlyGlnValLeuAspLeuThrValGluLysAspLys 17

145 ATAGCAACAGCAACAAATCTCATCATGCGCAATGCCACCAAGGAAAA 194

17 sLysAspLysArgGlyGlyGlnValAlaSerLysGlnSerGluGluGluProH 34

195 CGAAGAGATG.....GAGCAGCCCTCGAAGATGGACAGGAGACC 235

34 sHisLeuGluGluValGluAlaSerLysProGlyGlyAsn..... 47

236 GCCCTGTGGGAGGAGTGAGGGCCACCAGCCTCTCTCAACAACAACAAC 285

48ValArgArgLysValArgAr 54

286 AACACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 335

54 gLeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAspArg 71

336 ACTTGCCCTTAACCTCGATGGGCCATTCCTCCACAGGAGGATGATGACG 385

71 snGluGly.....GlyGluAspValGlyArgPheValValGlnGlyThr 85

386 GGTGGGTGGGAGATGGAGATGATGATAAATGTTGATGAGGAGATGAGA 435

86 GluValLysArgLysThrThrGlnGlnValArgProTyArgArgPhe 102

436 GAGATCGGGAGAAAGCTTAGGGAGCTACAGCTGAGAATTTGCTACGCAT 485

102 eArgThrProGluProAsnHisTyArgPheCysL 115

486 CCTATGGGGAGCTGCTTAACCAACCAACCAACCAACCAACCAACCA 535

115 euilePro 117

536 TTATGCT 543

seq_name: gb_sts:G72708

seq_documentation_block:

LOCUS G72708 421 bp DNA STS 08-AUG-2001

DEFINITION MARC 4953-4954:991939031:1 SCF - porcine spleen STS

genomic, sequence tagged site.

ACCESSION G72708

VERSION G72708.1 GI:15146738

KEYWORDS

STIS.

SOURCE

pig.

ORGANISM

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS

1 (bases 1 to 421)

Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L. and

Keele, J.W.

TITLE

Single nucleotide polymorphism (SNP) discovery in expressed porcine

genes

Unpublished (2001)

JOURNAL

COMMENT

Contact: Freking BA

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4278

Fax: 402 762 4173

Email: freking@email.marc.usda.gov

Primer A: GCAATGGGTGAACTCTACT

Primer B: CCAGCAGCAATAGACG

STS size: 500

PCR Profile:

Hotstart: 95 degrees for 15 minutes

Denature: 95 degrees for 30 seconds

Anneal: 56 degrees

Extension: 68 degrees for 2 minutes

Cycles: 32 to 45

Protocol:

Template: 50-200 ng genomic DNA

Primer: each 20 pmol

dNTPs: each 88 uM

Taq Polymerase: 0.25 units (Qiagen HotStar)

Buffer:

Commercially supplied Qiagen HotStar buffer

The STS is derived from PCR amplicons generated from genomic DNA,

sequenced from each end using the amplification primers. The


```

LOCUS       HSV351F8      45678 bp      DNA      23-NOV-1999
DEFINITION   Human DNA sequence from cosmid V351F8, between markers DXS366 and
              DXS87 on chromosome X contains ESTs.
ACCESSION    Z70719
VERSION      Z70719.1 GI:1261915
KEYWORDS     X.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
              Whiteley,M.
              Direct Submission
              Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
              CB10 1RQ, UK. E-mail enquiries: humquerry@sanger.ac.uk
              IMPORTANT: This sequence is the entire insert of clone V351F8. The
              true left end of clone V351F8 is at 1 in this sequence. The true
              right end of clone V351F8 is at 45678.
              V351F8 is from the human chromosome X-specific cosmid library.
FEATURES             Location/Qualifiers
     source          1..45678
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /map="X"
                     /clone="GHC-351F8"
                     /clone_lib="SCcv"
                     1..70
     repeat_region   1..45678
                     /partial
                     /note="Alu repeat: matches 80. .1 of consensus"
                     642..930
     repeat_region   1406..5179
                     /note="Alu repeat: matches 1. .308 of consensus"
     repeat_region   5243..5276
                     /note="L1 element fragment"
     repeat_region   5279..5567
                     /note="17 copies of 2 mer 82 % conserved"
                     /partial
                     /note="Alu repeat: matches 304. .1 of consensus"
                     9975..10369
     repeat_region   10399..12017
                     /note="MSTA element fragment"
     repeat_region   11318..111978
                     /note="MSTAR element fragment"
     repeat_region   11715..111846
                     /note="THE1B element fragment"
     repeat_region   12074..12170
                     /note="MLTIR element fragment"
     repeat_region   12227..12486
                     /note="MSTA element fragment"
     repeat_region   12437..12502
                     /note="MSTA element fragment"
     repeat_region   12756..13059
                     /note="MSTC element fragment"
                     /partial
                     /note="Alu repeat: matches 308. .1 of consensus"
                     14478..14650
     repeat_region   14491..14536
                     /note="THE1B element fragment"
     repeat_region   14715..14831
                     /note="MSTA element fragment"
     repeat_region   15099..15533
                     /note="THE1B element fragment"
     repeat_region   15534..15626
                     /note="L1 element fragment"
     repeat_region   15566..15619
                     /note="31 copies of 3 mer 85 % conserved"
     repeat_region   16352..16436
                     /note="3 copies of 18 mer 98 % conserved"
     repeat_region   16728..16763
                     /note="L1 element fragment"
     repeat_region   16764..17054
                     /note="9 copies of 4 mer 94 % conserved"

repeat_region       19420..19449
                     /partial
                     /note="Alu repeat: matches 308. .1 of consensus"
repeat_region       20071..20114
                     /note="15 copies of 2 mer 87 % conserved"
misc_feature        20071..20114
                     complement(19537..20013)
                     /note="match: 3' EST N51315 clone 283089"
repeat_region       20071..20114
                     /note="22 copies of 2 mer 98 % conserved"
repeat_region       20073..20112
                     /note="10 copies of 4 mer 100 % conserved"
repeat_region       20073..20126
                     /note="3 copies of 18 mer 87 % conserved"
repeat_region       23088..23300
                     /note="L1 element fragment"
repeat_region       23311..23490
                     /note="MLN1A element fragment"
repeat_region       23318..23489
                     /note="MLN1B element fragment"
repeat_region       23362..23491
                     /note="MLTIC element fragment"
repeat_region       23507..24407
                     /note="L1 element fragment"
repeat_region       24376..24462
                     /note="MSTC element fragment"
repeat_region       24379..24491
                     /note="MSTA element fragment"
repeat_region       24688..24738
                     /note="MSTA element fragment"
repeat_region       24810..25016
                     /note="L1 element fragment"
repeat_region       25254..25421
                     /note="L1 element fragment"
repeat_region       25439..25596
                     /note="MER25 element fragment"
repeat_region       26284..26343
                     /note="MLTID element fragment"
repeat_region       26428..26491
                     /note="MLTID element fragment"
repeat_region       26465..26518
                     /note="MLN1E element fragment"
repeat_region       27446..27667
                     /note="L1 element fragment"
repeat_region       28954..29247
                     /partial
                     /note="Alu repeat: matches 308. .1 of consensus"
                     30160..30325
repeat_region       32588..32875
                     /note="2 copies of 83 mer 98 % conserved"
misc_feature        32588..32875
                     /note="match: 5' EST H68599 clone 239077"
misc_feature        complement(32825..33230)
                     /note="match: 3' EST H68239 clone 289077"
repeat_region       35551..35802
                     /partial
                     /note="Alu repeat: matches 1. .260 of consensus"
                     35804..35839
repeat_region       38625..38705
                     /note="18 copies of 2 mer 83 % conserved"
repeat_region       38747..38839
                     /note="MLN1R element fragment"
repeat_region       38949..39032
                     /note="MLTIR element fragment"
repeat_region       38965..39076
                     /note="MLTIE element fragment"
repeat_region       38965..39076
                     /note="MLTID element fragment"
BASE COUNT      14260 a 9135 c 9327 g 12956 t
ORIGIN
alignment_scores:  Quality: 138.00      Length: 127
                   Ratio: 2.000         Gaps: 4
Percent Similarity: 54.331      Percent Identity: 31.496

```

alignment_block: US-09-327-750D-35 x HSV351F8

alignment_block: US-09-327-750D-35 x HSV351F8

Align seq 1/1 to: HSV351F8 from: 1 to: 45678

11 LeuThrValGluLysAspLysLysAspLysArgGlyGlyLysAlaSerLy 27
|||||:::||||| ||||| ||||| ||||| ::|||
10100 CCGCCGCCGGCAGGCCTGGTGGGCGGTGGGCGTCACAT 10457

27 sGlnSerGluGluGluProHisHisLeuGluGlu.....ValGluA 41

19458 GGAAATGTCCCCAGGAAAACAAAGTTGTGGACAAGCCCCAGTGCAGA 19507

41 SN.....LysLysProGlyGlyAsn 47

19508 ATGAAGCCCCCGCTTTAGGAGGTGGTGAATACCAGGAGCCTGGAGCAAT 19557

48 ValArgArgIysValArgArgLeuValProAsnPhcLeuTrpAlaIlePr 64

19558 GTTAAAGGGGTTTGGGCTCCACCTGCCCCGGGTTTGGAGAGGATGTGCC 19607

64 oAsnArgHisValAspArg.....AsnGluGluGluGluAspV 77

19608 CAATAGGCTTGTCCGATAACATTTGATATCATAGATGGAGATGGAGATGGATA 19657

77 a l g l v a r g p h e v a l v a l g l n g l y t h r g l u v a l l y s a r g l y s t h r t h r g l u 93

19658 TGGAACTGGAGGAGATGAGAGAGCTAAGGAGGAAATTAGGGAA 19707

94 GlnGlnValArgProTyrArgArgPheArgThrProGluPro..... 107

19708 CTTCAGTTGAGGTACAGTCTGCGCATTCCTATAGGGGACCCCTCACCACCA 19757

108 AGAAGNHSYVAsppheCysLeuTlepro 117

19758 TGATCATCATGATGAGTTTTGCCCTATGCCCT 19788

seq name: qb sts:G35294

seq_documentation_block:	477 bp	DNA	STS	02-OCT-1997
LOCUS	G35294			

LOCUS	DEFINITION	human	STS SHGC-37409,	477 bp	DNA	STS
		G35294				sequence tagged site.

ACCESSION
G35294.1
VERSION
G35294.1
GI:2459462

KEYWORDS str, str sequence, primer, sequence, sequence analysis
SOURCE human.

ORGANISM: none
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 477)

JOURNAL Unpublished (1997)

CONTACT: Richard M

Stanford University

Tel: 4157259687

Email: myers@shgc.

Primer A: AACATCTTTCCATGAAAGTTGATG

STS size: 106

Initial incubation: 95

Denaturation: 94

polymerization: 72

Thermal Cycler: per

Connor, R.

Connor, R.

AF097438.1 GI:4580589
house mouse.
Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Brown, A.L. and Kay, G.F.
Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 835)
Brown, A.L. and Kay, G.F.
Direct Submission
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
Location/Qualifiers
1. .835
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"
/map="near pip"
/tissue_type="blastocyst"
1. .835
/gene="Bex1"
205..591
/gene="Bex1"
/codon_start=1
/product="brain expressed x-linked protein 1"
/protein_id="A024429.1"
/db_xref="GI:4580590"
/translation="MESKQGVKNLNMENDHKKERKQDTRIRPEPAVALISAG
KNCAPGRGRRFRVQPIAHYRWDLRVGPQGRMREENVQREGDVRQMLEKLRER
QLSHSLRAVSTDPHPHHDHDFCLMP"
BASE COUNT 238 a 184 c 246 g 167 t
ORIGIN
alignment_scores:
Quality: 107.50 Length: 129
Ratio: 1.473 Gaps: 6
Percent Similarity: 56.589 Percent Identity: 30.233
alignment_block:
US-09-327-750D-35 x AF097438 ..
Align seg 1/1 to: AF097438 from: 1 to: 835
1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp... 16
205 ATGAGTCCAAAGATCAAGCGTGAATAATCTCAACATGGAGATGACCA 254
17LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGlu 32
255 TCAGAAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 304
32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
305 AGCAGCTGTGGCCCTGATCTCCGAGGCTGGCCAAAAGTGTCCGCTAGA 354
45 GlyGlyAsnValArgArgLysValArgLeuValProAsnPheLeuTr 61
355 GGAGGTGCGAGCGGTTCGGGGTTCGGAGCCATCGCTACTATAGATG 404
61 palaiIleProAsnArgHisValAspArgAsnGluGly.....GlyG 75
405 GGACCTGATCAGAGG...GTTGGGAGCCCGAGGAGGATGAGAGAGG 451
75 luAspValGlyArgPhePheValGlnGlyThrGluValLysArgLysThr 91
|||||

/note="MIR repeat: matches 102. .143 of consensus"
25378. .25555
/note="L2 repeat: matches 2330. .2519 of consensus"
25673. .25910
/note="MIR repeat: matches 4. .255 of consensus"
26722. .26798
/note="L2 repeat: matches 2572. .2649 of consensus"
27040. .27107
/note="L2 repeat: matches 2489. .2500 of consensus"
27108. .27374
/note="AluV repeat: matches 41. .304 of consensus"
27375. .27471
/note="L2 repeat: matches 2372. .2489 of consensus"
27651. .27945
/note="AluX repeat: matches 1. .299 of consensus"
28235. .28876
/note="L2 repeat: matches 1113. .1817 of consensus"
28975. .29097
/note="MER96 repeat: matches 42. .173 of consensus"
29173. .29196
/note="L2 copies 2 mer aa 100 conserved"
29832. .30260
/note="match: GSS: Em:AQ075670"
30392. .30597
/note="MIR repeat: matches 35. .252 of consensus"
30845. .31010
/note="MER91B repeat: matches 2. .162 of consensus"
31056. .31094
/note="MIR repeat: matches 218. .256 of consensus"
31728. .31791
/note="L6 copies 4 mer ttgtg 76 conserved"
31858. .31952
/note="MIR repeat: matches 48. .153 of consensus"
complement(3437. .33589)
/note="match: STS: Em:H55731"
33543. .33633
/note="MIR repeat: matches 84. .165 of consensus"
33785. .33966
/note="MER91A repeat: matches 1. .185 of consensus"
alignment_scores:
Quality: 114.50 Length: 66
Ratio: 2.602 Gaps: 3
Percent Similarity: 66.667 Percent Identity: 39.394
alignment_block:
US-09-327-750D-35 x HS714B7 ..
Align seg 1/1 to: HS714B7 from: 1 to: 98274
55 LeuValProAsnPheLeuTrpAlaIleProAsnArgHisValAsp..... 69
62887 TTGCCCTAATTTTCAATGGCCATACCAATAGCAGGTCATGATGG 62936
70 .ArgAsnGluGlyGlyGluAspValGlyArgPheValValGlnGlyThrG 86
62937 GATCAATAGAGTGGAGATGATGGAATGTTTCATGGAGGATGAGAG 62986
86 luValLysArgLys.....ThrThrGluGlnGlnValArgPro 98
62987 GAATCAGAGAGATAGGAGCTACAAATTGAGGAATTTCTGTGTATCCT 63036
99 TyrArgArgPheArgThrProGluProAsnHisTyrAspPheCys 114
63037 TATGGAGCTGTGTAATCCCAT...GACCATCATGATGAATTTTGC 63081
seq_name: gb_ro:AF097438
seq_documentation_block:
LOCUS AF097438 835 bp mRNA ROD 11-APR-1999
DEFINITION Mus musculus brain expressed x-linked protein 1 (Bex1) mRNA,
complete cds.
ACCESSION AF097438

us-09-327-750d-35.rge

Tue Mar 12 09:01:45 2002

452 AGAACGTACAGAGTTTGGGGTGATGTGACACAGCTCATCGAGAAGCTG 501
 92 ThrGluGlnGlnValArgProTyrArgPheArgThrProGluPro.. 107
 502 AGGGAAGGACGCTGAGCCAGAGCTGGGGCGGTAGCACTGACCCGCC 551
 108AspAsnHisTyrAspPheCysLeuLeuPro 117
 552 TCATCATGACCACTGATGAGTTTGGCTCATGCC 588
 seq_name: gb_ro:AF051347
 seq_documentation_block: 838 bp mRNA ROD 01-OCT-1998
 LOCUS AF051347
 DEFINITION Mus musculus REX-3 mRNA, complete cds.
 ACCESSION AF051347
 VERSION AF051347.1 GI:3510642
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 838)
 Faria,T.N., LaRosa,G., Wilen,E., Liao,J. and Gudas,L.J.
 Characterization of genes which exhibit reduced expression during
 the retinoic acid-induced differentiation of F9 teratocarcinoma
 cells: involvement of cyclin D3 in RA-mediated growth arrest
 Mol. Cell. Endocrinol. 143 (1-2), 155-166 (1998)
 99021197
 2 (bases 1 to 838)
 Faria,T.N., LaRosa,G., Wilen,E., Liao,J. and Gudas,L.J.
 Direct Submission
 Submitted (26-FEB-1998) Pharmacology, Cornell University Medical
 College, 1300 York Avenue, New York, NY 10021, USA
 Location/Qualifiers
 1..838
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /cell_line="F9 teratocarcinoma"
 229..579
 /note="retinoic acid reduced gene expression"
 /codon_start=1
 /product="REX-3"
 /protein_id="AAC61929.1"
 /db_xref="GI:3510643"
 /translation="MENDHOKKEEKEKPODITRREPAVALISEAGKNCAPRGRRRF
 RVROPFAHYRWDLQMVGEFGQRMREENVQFGDVRQLMEXLRQLRSLRAYSTD
 PPHDHDFECLMP"
 BASE COUNT 238 a 181 c 251 g 168 t
 ORIGIN

alignment_scores:
 Quality: 107.50 Length: 129
 Ratio: 1.473 Gaps: 6
 Percent Similarity: 56.589 Percent Identity: 30.233
 alignment_block:
 US-09-327-750D-35 x AF051347 ..
 Align seg 1/1 to: AF051347 from: 1 to: 838
 1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp.. 16
 193 ATGGAGTCCAAAGATCAAGCGTGAAGAAATCTCAACATGGAGAATGACCA 242
 17LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluG 32
 243 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 292
 32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
 293 AGCCAGCTGTGCGCTCATCTCCGAGGCTGGCAAAACCTGTGCGCCTAGA 342

alignment_block:
 US-09-327-750D-35 x AF051347 ..
 Align seg 1/1 to: AF051347 from: 1 to: 838

1 MetAlaSerLysLysGlnValIleLeuAspLeuThrValGluLysAsp.. 16
 193 ATGGAGTCCAAAGATCAAGCGTGAAGAAATCTCAACATGGAGAATGACCA 242
 17LysLysAspLysArgGlyGlyLysAlaSerLysGlnSerGluGluG 32
 243 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 292
 32 luProHisHisLeu.....GluGluValGluAsnLysLysPro... 44
 293 AGCCAGCTGTGCGCTCATCTCCGAGGCTGGCAAAACCTGTGCGCCTAGA 342

45 GlyGlyAsnValArgArgLysValArgLysValProAsnPheLeuTr 61
 343 GGAGGTGCGAGCGGTTCGGGTTCGGAGCCCATCGCTCCTACTATAGATG 392
 61 pAlaIleProAsnArgHisValAspArgAsnGluGly.....GlyG 75
 393 GGACCTGATGCAGAGG...GTTGGGAGCCCGCAGGAAGGATGAGAGAGG 439
 75 luAspValGlyArgPheValValGlnGlyThrGluValLysArgLysThr 91
 440 AGAACGTACAGAGTTTGGGGTGATGTGACACAGCTCATCGAGAAGCTG 489
 92 ThrGluGlnGlnValArgProTyrArgPheArgThrProGluPro.. 107
 490 AGGAAAAGCTAGCTGAGCCACAGCCCTGGGGCGGTAGCACTGACCCGCC 539
 108AspAsnHisTyrAspPheCysLeuLeuPro 117
 540 TCATCATGACCACTGATGAGTTTGGCTCATGCC 576


```
;
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 326..6277
; US-08-836-325-9

alignment_scores:
  Quality: 72.50 Length: 58
  Ratio: 1.812 Gaps: 4
  Percent Similarity: 68.966 Percent Identity: 36.207

alignment_block:
  US-09-327-750D-34 x US-08-836-325-9/rev ..
  Align seg 1/1 to reverse of: US-08-836-325-9 from: 1 to: 6452
    25 LysAlaSerLysGlnSerGluGluGlu.....SerHisHisLeuG1 38
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
    3220 AAGAGCCAGAAACAGAGTTCCAGGACCAACAGGTTCCCAATCACCATGACCA 3171
    38 uGluValGluAsnLysLysProGlyGlyAsnValargLysValargA 55
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    3170 TCATGTAACAATAAGACACATGGT.....CTGCCCGCCGACCCPC...C 3130
    55 rgLeuValProAsnPhLeuTrpAlaIleProAsnArgHisValasPhis 71
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    3129 ATGCAGTCCCAATGGTCTCTATCCACTCCCAACACAGCAGCTCGGAACAC 3080
    72 SerGluGlyGly....GluGluVal 78
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    3079 GATGAGGAAGAGTGGGAAGTC 3056

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-892-770-1

seq_documentation_block:
; Sequence 1, Application US/08892770
; Patent No. 5962670
; GENERAL INFORMATION:
; APPLICANT: Walling, Linda L.
; APPLICANT: Pautot, Veronique
; APPLICANT: Gu, Yong-Qiang
; APPLICANT: Chao, Wun Shaw
; TITLE OF INVENTION: Improved Promoters for Enhancing Plant
; TITLE OF INVENTION: Productivity
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,770
; FILING DATE: 15-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Quine, Jonathan A.
; REGISTRATION NUMBER: P-41,261
; REFERENCE/DOCKET NUMBER: 023070-072100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 935 base pairs

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY:
; LOCATION: 1..935
; OTHER INFORMATION: /note= "tomato acidic leucine
; OTHER INFORMATION: aminopeptidase 1 (LapAl) promoter
; OTHER INFORMATION: region"
; US-08-892-770-1

alignment_scores:
  Quality: 72.00 Length: 60
  Ratio: 2.250 Gaps: 2
  Percent Similarity: 53.333 Percent Identity: 28.333

alignment_block:
  US-09-327-750D-34 x US-08-892-770-1 ..
  Align seg 1/1 to: US-08-892-770-1 from: 1 to: 935
    62 TrpAlaIleProAsnArgHisValasPhisSerGluGlyGlyGluValA 78
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    438 TGGTGTGGTGAATGAGATTATCATGATCAATAAGAAAGAGGGGAGAAGAT 487
    78 lclYargPheValGlyGlnValMet.GluAlaLysArgHis..... 91
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    488 GAACAGTGTTTTAAATAAATTTTGGCCCAAGAAACATTCACCGCT 537
    92 .....SerLy 93
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    538 GACITTAAGTCATTACATCACTTCCTGCGATGACAGAGAAATCGTCAAA 587
    93 sGluGlnMetArgProTyrThrArg 102
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
    588 AAATGACACAATGAGACCTACATAAGG 615

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-726-214-5

seq_documentation_block:
; Sequence 5, Application US/08726214
; Patent No. 6107076
; GENERAL INFORMATION:
; APPLICANT: Tang, Wei-Jen
; APPLICANT: Gilman, Alfred G.
; TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,214
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,498
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:450
```



```

Percent similarity: 59.000      Percent Identity: 20.000
Alignment block:
US-09-327-750D-34 x 5180810-3  ..

Align seg 1/1 to: 5180810-3 from: 1 to: 1131

3  SerIysValLysGlnValIleLeuAspLeuThrvAlGluLysAspLysLys 19
178 GCTAAACTGTTGCAGGTTTGAACACACCAAGTTTGGAAACGAA..... 222
19  sAsnLysLysGlyLysAlaSerLysGlnSerLysGluGluSerHisH 36
223 .....AAACTCAAGAGTGAGAAATGAGAGAATAAGAAAA 256
36  lSteuGluGluValGluAsnLysLysProGlyGlyAsnValAlGAgLys 52
257 ATTTAGACAACCTTAGCAAGAAATAATCAAGA.....AAG 291
53  ValArgArgLeuValProAsnPheLeuTrpAlaIleProAsnArgHisVa 69
292 CTCGAAAAATTTGGAGCTTGACTATCTCAAAAAATTAGATCACGAGACAA 341
69  lAspHisSerGluGlyGlyGluGluValGlyArgPheValGlyGlnValM 86
342 AGAGCCCAAAAAAACAACACAGAA.....C 367
86  etGluAlaLysArgHisSerLysGluGlnMetArgProTyrThrArg 102
368 AAGAAGAGCGCACAAAAAATCAAGAACCAATTAGAACCGTAAATACCAACGA 417

```

```

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:09-045-583-6
seq_documentation_block:
: Sequence 6, Application US/09045583
: Patent No. 6287805
: GENERAL INFORMATION:
: APPLICANT: Graham, Gerard J. et al.
: TITLE OF INVENTION: No. 6287805el Molecules of the G Protein-Coupled
: NUMBER OF SEQUENCES: 56
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD, LLP
: STREET: 28 State Street
: CITY: Boston
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/045,583
: FILING DATE: 20-MAR-98
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Mandragoras, Amy E.
: REGISTRATION NUMBER: 36,207
: REFERENCE/DOCKET NUMBER: MNT-044
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617)227-7400
: TELEFAX: (617)742-4214
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1137 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4533 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-726-214-5
;
alignment_scores:
    Quality: 72.00      Length: 117
    Ratio: 1.075        Gaps: 5
Percent Similarity: 57.265 Percent Identity: 26.496
alignment_block:
US-09-327-750D-34 x US-08-726-214-5 ..
Align seg 1/1 to: US-08-726-214-5 from: 1 to: 4533
2 AlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAspLy 18
||||| : : : : : ||||| : : : : :
1918 GCATCAAGCCACCTCCCCTGCCCTATTGAGACAACAGGAGCCCAATGG 1967
||||| : : : : : ||||| : : : : :
18 stlyasnlyslysGlyGlyLysAlaSerLysGlnSerGluGluSerH 35
::: : : : : : : : : : : : : : : : : : : : : : : : : : :
1968 GAGTGCCTCATGCAGCGCTCCACATCAGAGGAGGCTGAAGAA..... 2010
||||| : : : : : ||||| : : : : :
35 iShisLeuGluCluValGluasnlysLysProGly...GlyAsnValArg 50
::: : : : : : : : : : : : : : : : : : : : : : : : : : :
2011 .....CAGGAGGCCCGAGGCTGACAAACCCTCGTCCCCAACCCCGC 2052
||||| : : : : : ||||| : : : : :
51 ArgLysValArgLeuValProAsnPheLeuTrpAlaIleProAsnAr 67
||||| : : : : : ||||| : : : : :
2053 CGCAGGCTCGC.....CTCCAGGACCTGGCAGACCG 2084
||||| : : : : : ||||| : : : : :
67 qHisValAspHisSerGluGlyGlyGluGluValGlyArgPheValGly 84
::: : : : : : : : : : : : : : : : : : : : : : : : : : :
2085 TGTGTTGGAGCGCTCTGAGGATGAGCACGAACTGAACACGACTTCTTAACG 2134
||||| : : : : : ||||| : : : : :
84 lnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArgProTy 100
::: : : : : : : : : : : : : : : : : : : : : : : : : : :
2135 AGGCCCTG...CTGAGCGGAGTGTCGCCCGAGGTGTTAAGRAGAANAAC 2181
||||| : : : : : ||||| : : : : :
101 Thr.....ArgPheArgThrProGluProAspAsnHisty 112
||||| : : : : : ||||| : : : : :
2182 ACATTCTCTTAACGATGAGTTCATGACCACCGAGATGGAACACGCTA 2231
||||| : : : : : ||||| : : : : :
112 r 112
2232 C 2232
seq_name: /cgn2_6/ptodate/2/ina/backfiles1.seq:5180810-3
seq_documentation_block:
; Patent No. 5180810
; APPLICANT: Gomi, Hideyuki; Hozumi, Tatsunobu; Hattori, Shizuo;
; Tagawa, Chikaki; Kishimoto, Fumitaka; Bjorck, Lars
; TITLE OF INVENTION: PROTEIN H CAPABLE OF BINDING TO IGG
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/376,641
; FILING DATE: 07-JUL-1989
; SEQ ID NO:3
; LENGTH: 1131
5180810-3
alignment_scores:
    Quality: 71.00      Length: 100
    Ratio: 1.203        Gaps: 3

```



```

alignment_scores:
  Quality: 71.00      Length: 100
  Ratio: 1.203       Gaps: 3
  Percent Similarity: 59.000   Percent Identity: 20.000

alignment_block:
  US-09-327-750D-34 x 5180810-2 ..

  Align seg 1/1 to: 5180810-2 from: 1 to: 1784

      3 SerLySvAllySGlnValIleLeuAspIeuThrValGluLyAspLyLy 19
      :|||||: :|||||: :|||||:
      505 GCTAAACTGCTGAGCTTGTCGAACCAACGAGTTGGAAACGAA..... 549
      19 SASnLyLySGLyGlyLyAlaSerLySGInSerGluGluGlnGluSerHisH 36
      :|||||: :|||||: :|||||:
      550 ..... AACCTCAAGAGCGAAGATAGAGGAGATAGAA 583
      36 ISteuGluGluValGluAsnLyLySProGlyGlnValAlaArgArgLyS 52
      :|||||: :|||||: :|||||:
      584 ATTAGACAAACTTAGCAAGAAATTCAGAA.....AAG 618
      53 ValArgArgLeuValProAsnIleuThrPalalIleProAsnArgHisVa 69
      :|||||: :|||||: :|||||:
      619 CTCGAAAATTGGAGCTGACTATCTCAAAAAATTAGATCAGACAGACAA 668
      69 IAspHisSerGluGlyGlyGluGluValGlyArgPheValGlyGlnValM 86
      :|||||: :|||||: :|||||:
      669 AGAGACCCCAAAACCAACCAAGAA.....C 694
      86 eEtGAlAlaLySARGHISerLySGInGluGlnMetArgProTyrrThArg 102
      :|||||: :|||||: :|||||:
      699 AAGAGAGCGGCAAAAAAATCAAGAACATTAGAAGCTAATATCCAAACA 744

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-841-483-3

seq_documentation_block:
  Sequence 3, Application US/08841483B
  Patent NO. 5978875
  GENERAL INFORMATION:
  APPLICANT: Prescott, Steven M.
  APPLICANT: Bunting, Michaelne
  APPLICANT: Tang, Wen
  APPLICANT: Topham, Matthew
  TITLE OF INVENTION: Discylglycerol Kinase Isoforms Epsilon and Zeta and
  TITLE OF INVENTION: Methods of Use Thereof
  FILE REFERENCE: 2037.2.1a
  CURRENT APPLICATION NUMBER: US/08/841.483B
  CURRENT FILING DATE: 1997-04-22
  EARLIER APPLICATION NUMBER: 60/016,210
  EARLIER FILING DATE: 1996-04-22
  NUMBER OF SEQ ID NOS: 33
  SOFTWARE: PatentIn Ver. 2.0
  SEQ ID NO 3
  TYPE: DNA
  LENGTH: 3490
  ORGANISM: Homo sapiens
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (89)..(2875)
  US-08-841-483-3

alignment_scores:
  Quality: 71.00      Length: 82
  Ratio: 1.651       Gaps: 4
  Percent Similarity: 52.459   Percent Identity: 30.488

alignment_block:
  US-09-327-750D-34 x US-08-841-483-3 ..

  Align seg 1/1 to: US-08-841-483-3 from: 1 to: 3490

```

```

13  ThValcIuLbysAspIlystIysAsnIlystGlyGly.....LysAl 26
848  ACTCTGAACCAACGACCAAGAGAGAGAGAGAGCGCATCTTTCANAGSAAATC 897
26  aSerIyGlnSerGlInuGlInuSerHis..... 35
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
898  CAGCAGAAAGAGGGCGCTGAGAGAGGGCGCGCTGGAGACCCCTTCATCATCAGC 947
36  .....HisLeuGlInuIValcIuAsnIlystIys 44
      ||||| |||| |||| |||| |||| |||| |||| |||| ||||
948  CCACCCCTCCCGCCTCATGAAACCCCTGCTGTGTGTTGTGTAAGACCCCAAG 997
45  ProGlIyGlnsValArgArgIyValArgArgLeuValProAsnPhle 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
998  AGTGGGGGACCAACGAGGTGCAGAAAGATC.....ATCCAGTCTTTTCT 1038
61  uTrpAlaIleProAsnArgHisVal...AspHisSerGlInuIyGly 75
      ||||| :: ||||:|||| ||||:|||| ||||:|||| ||||:||||
1039  CTGGATCTCATATCCGACCAAGACTCTTCGACCTGACCCAGCCAGGGAGGG 1084
seq_name: /cgn2_6/prodata/2/lna/6B_COMB.seq:US-09-382-911-3

```

```

seq documentation_block:
Sequence 3, Application US/09382911
Patent No. 6221658
GENERAL INFORMATION:
APPLICANT: Prescott, Steven M.
APPLICANT: Bunting, Michaeline
APPLICANT: Tang, Wen
APPLICANT: Topham, Matthew
TITLE OF INVENTION: Diacylglycerol Kinase Isoforms Epsilon, and Zeta and
FILE OF INVENTION: Methods of Use Thereof
FILE REFERENCE: 2037.2.1a
CURRENT APPLICATION NUMBER: US/09/382,911
CURRENT FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 08/841,483
PRIOR FILING DATE: 1997-04-22
PRIOR APPLICATION NUMBER: 60/016,210
PRIOR FILING DATE: 1996-04-22
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 3490
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (89)..(2875)
US-09-382-911-3

alignment_scores:
Quality: 71.00 Length: 82
Ratio: 1.651 Gaps: 4
Percent Similarity: 52.439 Percent Identity: 30.488

alignment_block:
US-09-327-750D-34 x US-09-382-911-3 ..
Align seg 1/1 to: US-09-382-911-3 from: 1 to: 3490

```

13 ThrValGluLysAspLysLysAsnLysLysGlyGly.....LysAl 26
848 ACCTGTGAACCAACGACAAAGAGAGAGAGGCGATCTCTCAGAGAGAGTGC 897
26 aserLysGlnserGluGluGlnserHis..... 35
898 CACGACAGAAAGGGCGCTGAGAGAGGGCGCGCTGAGACCTTCATCATCAGGC 947
36HisLeuGluGluValGluAsnLysLys 44
948 CCAACCCCTCCCGCTCATGAAACCCCTCTGAGTGTGTGTGAACCCACAG 997

COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/807,043B
FILING DATE: 19911212
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/764,364
FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,838
FILING DATE: 9-JULY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705,702
FILING DATE: 23-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 534274man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 253.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4698 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: singular
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-07-807-043B-5

alignment_scores:
Quality: 71.00 Length: 93
Ratio: 1.365 Gaps: 4
Percent Similarity: 55.914 Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-07-807-043B-5/rev ..

Align seg 1/1 to reverse of: US-07-807-043B-5 from: 1 to: 4698

```
16 LysAspLysLysAsnLysGlyLysAlaSer LysGlnSerGlnG 32
:::|||||:|||||:|||||:|||||:|||||:|||||:
2134 AGGACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 2085
32 LngLysHisHisLysGlnGluValGluAsnLys.....Lys 44
:::|||||:|||||:|||||:|||||:|||||:|||||
2084 AGGAAAGGCACACAGAGGAGGAGGAGGAGGAGGAGGAGGATCTTGGAAA 2035
45 ProGlyLysAsnValArgIleValArgIleValArgIleValProAsnPhenLe 61
|||||:|||||:|||||:|||||:|||||:|||||:
2034 CAAAGTGCACAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1990
61 UTPrAlaIleProAsnArgHisValAspHisSerGlyGlyGlnGluVal 78
|||||:|||||:|||||:|||||:|||||:|||||
1989 .....CCAAACAAACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1950
78 aLgLy.....ArgPheValGlyGlnValMetGluAlaLys 89
|||:|||||:|||||:|||||:|||||:|||||
1949 CAGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1900
90 ArgHisSerLysGlnGlnGlnMetArg 98
:::|||||:|||||:|||||:|||||:|||||:|||||
1899 AAGGCTGCTAAAGTCGAGCAAGAGAGC 1873
seq_name: /cgn2_6/ptodata/2/lna/5A_Comp.seq:US-08-299-849B-5
seq documentation_block:
; Sequence 5, Application US/08299849B
; Patent No. 5612201
; GENERAL INFORMATION:
```

APPLICANT: De Plaen, Etienne; Boon-Falleur, Thierry;
APPLICANT: Leth, Bernard; Szikora, Jean-Pierre; De Smet, Charles;
APPLICANT: Chomez, Patrick
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful In
TITLE OF INVENTION: Determining Expression Of A Tumor Antigen Precursor
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/299,849B
FILING DATE: 1-SEPTEMBER-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/037,230
FILING DATE: 26-MARCH-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04354
FILING DATE: 22-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/807,043
FILING DATE: 12-DECEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/764,364
FILING DATE: 23-SEPTEMBER-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,838
FILING DATE: 9-JULY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/705,702
FILING DATE: 23-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5612201man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5355
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4698 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-299-849B-5

alignment_scores:
Quality: 71.00 Length: 93
Ratio: 1.365 Gaps: 4
Percent Similarity: 55.914 Percent Identity: 27.957

alignment_block:
US-09-327-750D-34 x US-08-299-849B-5/rev ..

Align seg 1/1 to reverse of: US-08-299-849B-5 from: 1 to: 4698

```
16 LysAspLysLysAsnLysGlyLysAlaSer LysGlnSerGlnG 32
:::|||||:|||||:|||||:|||||:|||||:|||||:
2134 AGGACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2085
32 LngLysHisHisLysGlnGluValGluAsnLys.....Lys 44
:::|||||:|||||:|||||:|||||:|||||:|||||
2084 AGGAAAGGCACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGATCTTGGAAA 2035
```


alignment_scores:

Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:
 US-09-327-750D-34 x AAH03517

Align seg 1/1 to: AAH03517 from: 1 to: 865

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 ATGGAGTCCAAAGAGAACTAGCGCAACAACTCTCAACGGGGAATAATGC 240
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGlu 33
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 CCACAAAGAAACGAAGAGGAGGAGAGCCGCCACGACGATGAGAGAG 290
33 luseRHisHisLeuGluGluValGluAsnLysProGluGlyAsnVal 49
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 AATCCCGCATTTGGAGGGGCTGAGAGCCAGAAAGCTTGAGGAAATATC 340
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIle 65
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 AGCGGGGGGCGAGTTAGCGACTTCTCCCTAATTTTCATGGCGCATACC 390
65 oAsnArgHisValAspHisSerGluGluGluValGlyArgPheV 82
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 TAAATAGCATATTGACACAAATGAGAGCGAGATGATGTAAGAGGTTG 440
82 alGlyGlnValMetGluAlaLysArgHisSerLysGlnGlnMetArg 98
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
441 TAGGGCAGATGATGCAAAATCAAGAGAAAGACTAGGAAACAGAGATGAG 490
99 ProTyrThrArgPheArgThrProGluProAsnHisTyrAspPheCy 115
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
491 CACTATATGCGCTTCCAAACTCTGTAACCTGACACCACTATATGACTTTG 540
115 sLeuIlePro 118
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
541 CCTCATATCCT 550

seq_name: /SID2/gcdata/geneseq/geneseqn/NA2001.DAT:AAH13750
seq_documentation_block:
ID AAH13750 standard; cDNA: 1229 BP.
XX
XX AAH13750;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human cDNA sequence SEQ ID NO:10656.
DE
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
KM
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
PN
XX
XX 07-FEB-2001.
PD
XX
XX 28-JUL-2000; 2000EP-0116126.
PF
XX
XX 29-JUL-1999; 99JP-0248036.
PR
XX 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241859.
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

```

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

PS Claim 8; SEQ ID 10656; 2537pp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH03446 to
 CC AAH05893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

SO Sequence 1229 BP; 351 A; 243 C; 296 G; 339 T; 0 other;

alignment_scores:

Quality: 372.00 Length: 120
 Ratio: 3.758 Gaps: 2
 Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:
 US-09-327-750D-34 x AAH13750

Align seg 1/1 to: AAH13750 from: 1 to: 1229

```

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
191 ATGGAGTCCAAAGAGAACTAGCGCAACAACTCTCAACGGGGAATAATGC 240
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGlu 33
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 CCACAAAGAAACGAAGAGGAGGAGAGCCGCCACGACGATGAGAGAG 290
33 luseRHisHisLeuGluGluValGluAsnLysProGluGlyAsnVal 49
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 AATCCCGCATTTGGAGGGGCTGAGAGCCAGAAAGCTTGAGGAAATATC 340
50 ArgArg...LysValArgArgLeuValProAsnPheLeuTrpAlaIle 65
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 AGCGGGGGGCGAGTTAGCGACTTCTCCCTAATTTTCATGGCGCATACC 390
65 oAsnArgHisValAspHisSerGluGluGluValGlyArgPheV 82
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 TAAATAGCATATTGACACAAATGAGAGCGAGATGATGTAAGAGGTTG 440
82 alGlyGlnValMetGluAlaLysArgHisSerLysGlnGlnMetArg 98
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
441 TAGGGCAGATGATGCAAAATCAAGAGAAAGACTAGGAAACAGAGATGAG 490
99 ProTyrThrArgPheArgThrProGluProAsnHisTyrAspPheCy 115
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
491 CACTATATGCGCTTCCAAACTCTGTAACCTGACACCACTATATGACTTTG 540

```



```
115 sleutlePro 118
|||||
541 CCTCATACCT 550

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAC85548
seq_documentation_block:
ID AAC85548 standard; cDNA; 1364 BP.
XX AC AAC85548;
XX DT 04-JUN-2001 (first entry)
XX DE cDNA encoding CDIFF-4, Incyte ID No. 1990956CBL.
XX KW Human; cell differentiation; CDIFF; agonist; antagonist; epilepsy;
XX KW cell proliferation; Alzheimer's disease; schizophrenic disorder;
XX KW arteriosclerosis; cancer; atherosclerosis; diabetes mellitus; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 250..612
XX FT /*tag= a
XX FT /product= "CDIFF-4"
XX PN WO200119860-A2.
XX PD 22-MAR-2001.
XX PF 14-SEP-2000; 2000WO-US25435.
XX PR 15-SEP-1999; 99US-0154140.
XX PR 06-DEC-1999; 99US-0169155.
XX PA (INCYTE) INCYTE GENOMICS INC.
XX PI Tang YT, Hillman JL, Yue H, Reddy R, Lal P, Shah P, Azimzai Y;
XX PI Baughn MR, Lu DM, Bandman O, Shih LL, Patterson C;
XX DR P-PSDB; AAB47126.
XX PT Isolated polypeptides and polynucleotides involved in cell
XX PT differentiation are used for treatment, prevention and diagnosis of
XX PT cell proliferative, developmental and neurological disorders e.g.
XX PT cancer and Alzheimer's disease.
XX PS Claim 5; Page 121; 137pp; English.
XX CC The sequences given in AAC85545-72 encode human polypeptides involved
XX CC in cell differentiation (CDIFF). CDIFF polypeptides and agonists of
XX CC these are used to treat a disease or condition associated with
XX CC decreased expression of functional CDIFF. An antagonist of CDIFF is
XX CC used to treat a disease or condition associated with over expression
XX CC of functional CDIFF. CDIFF polypeptides may be used for the treatment,
XX CC prevention and diagnosis of cell proliferative, developmental and
XX CC neurological disorders, such as Alzheimer's disease, schizophrenic
XX CC disorders, arteriosclerosis, cancer, atherosclerosis, diabetes mellitus
XX CC and epilepsy. The CDIFF-4 sequence is homologous to Mus musculus
XX CC REX-3. This sequence maps to chromosome 1 within the interval from
XX CC 152.2 to 157.4 centimorgans, to chromosome 3 within the interval from
XX CC 157.4 to 158.0 centimorgans, and to the X chromosome within the interval
XX CC from 104.9 to 150.3 centimorgans.
XX SQ Sequence 1364 BP; 411 A; 269 C; 322 G; 362 T; 0 other;

alignment_scores:
Quality: 372.00 Length: 120
Ratio: 3.758 Gaps: 2
Percent Similarity: 82.500 Percent Identity: 60.000
```

```
alignment_block:
US-09-327-750D-34 x AAC85548 ..
Align seg 1/1 to: AAC85548 from: 1 to: 1364
1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLysAs 17
||| ||||| : : : : : ||||| : : : : : |||||
250 ATGGAGTCCAAAGAGAACTAGCGCAAACTCTCAACGGGAAATGC 299
17 pLysLysAsnLysLysGlyGly...LysAlaSerLysGlnSerGluGluG 33
: : : : : : : : : : : : : : : : : : : : : : : : : : :
300 CCAACAAGAAACGAAGAGGAGGAGCGAGCCGCCACGACAGATGAAGAAG 349
33 LuSerHisLysLeuGluValGluAsnLysLysProGlyGlyAsnVal 49
||| ||||| : : : : : ||||| : : : : : |||||
350 AATCCCGCCATTGGGAGGGGTGAAGCCAGAAAGCCTGGAGGAATATC 399
50 ArgArg...LysValArgLeuValProAsnPheLeuTrpAlaIlePr 65
||| ||||| : : : : : ||||| : : : : : |||||
400 AGCGGGGCGAGTGTAGCGAGTTGTCCTAATTTGCGATGGCCATACC 449
65 OAsnArgHisValAspHisSerGluGlyGlyValGluValGlyArgPheV 82
||| ||||| : : : : : ||||| : : : : : |||||
450 TAATAGCATATTGACCAATGAAGCGAGAGATGATAGAAAGGTTTG 499
82 alGlyGlnValMetGluAlaLysArgHisSerLysGluGlnGlnMetArg 98
||| ||||| : : : : : ||||| : : : : : |||||
500 TAGGCGAGATGATGGAATCAAGAGAAAGACTAGGGAACAGAGATGAGG 549
99 ProTyrThrArgPheArgThrProGluProAspAsnHisTyrAspPheCy 115
||| ||||| : : : : : ||||| : : : : : |||||
550 CACTATATGCGCTTCCAAACTCTGACCTGACCAACCATTTGACTTTTG 599
115 sleutlePro 118
|||||
600 CCTCATACCT 609

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: AAF23529
seq_documentation_block:
ID AAF23529 standard; DNA; 891 BP.
XX AC AAF23529;
XX DT 22-MAR-2001 (first entry)
XX DE Human NADE DNA.
XX KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
XX KW neurogenetic disease; NF-kappaB; ds.
XX OS Homo sapiens.
XX PN WO200075278-A2.
XX PD 14-DEC-2000.
XX PF 07-JUN-2000; 2000WO-US15621.
XX PF 07-JUN-1999; 99US-0327750.
XX PR (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Sato T;
XX WPI; 2001-061707/07.
XX PT New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX PT the gene encoding NADE, useful for modulating the activity of p75NTR.
XX PT and for detecting neurodegenerative diseases.
XX PS Disclosure: Fig 1; 134pp; English.
```

XX The present invention relates to a purified polypeptide capable of
 CC binding neurotrophin receptor (p75-NTR). The invention is useful for
 CC binding and modulating the activity of p75NTR. The peptide mediates
 CC NGF-induced apoptosis, which plays an important role in neurogenetic
 CC diseases. The peptide of the invention and p75NTR are useful for
 CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
 CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
 CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.
 SO Sequence 891 BP; 251 A; 182 C; 224 G; 234 T; 0 other:

alignment_scores: Length: 126
 Quality: 171.50 Gaps: 3
 Ratio: 2.257
 Percent Similarity: 60.317 Percent Identity: 32.540

alignment_block:

US-09-327-750D-34 x AAF23529 ..

Align seg 1/1 to: AAF23529 from: 1 to: 891

```

12 LeuThValGLuLysAspLysAsnLysLysGlyLysAlaSerly 28
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 TTAAGTGTGTCAAGAAACACGAAACCAAAATCTCATGCGC 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
28 sGlnSerGluGluSerHisLsleuGluVal..... 40
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
317 AAATATTACACAGAAACAGAGATGAGACACCTATGAGATGAG 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 .....GluAsnLysProGlyGlyAsn 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
367 AGAAGACCCCTTGGAGAGAGTGAAGCCACGCTGACGGAAT 416
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
49 ValArgGlyValArgGlyValProAsnPhelThrPalalLepr 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 CGAGGGGAGAGGCTCGGACTTCCCTAAATTTGATGAGGCAATAC 466
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 oAsnArgHisValAspHisSerGly.....GlyGluGluValGly 80
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
467 CAATAGGAGATCAATGATGAGATGAGATGATGATGATGATAA 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 rGheValGlyValMetGluAlaLysArgHisSerlyGluGlu 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
517 TATTCATGAGAGATGAGAAATCAGAAACCTTATGAGAGCTCAG 566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 MetArgProGlyThrArgPheArgThrProGluProAsnHisTyrAs 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
567 TTGAGGAATGTCTGCTATCTTATGAGGAGCTCTTATATCACAATGA 616
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
113 P.....PheCysLeuIlePro 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
617 CCATCATGATGATTTTGCCTTATGCT 644

```

seq_name: /stids/gcgdata/geneseq/geneseq/NA2000.DAT:AAF21748

seq_documentation_block:

ID AAF21748 standard; DNA; 917 BP.

XX AAF21748;
 AC
 XX
 DT 27-MAR-2001 (first entry)
 DE Human breast and ovarian cancer associated antigen gene SEQ ID 135.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neurotrophic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antibacterial; antiinflammatory; anticancer; vulnerrary; anticonvulsant;
 KW addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;

KW cardiovascular disorder; wound healing; neurological disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200055173-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000MO-US05881.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Rubin SM;
 XX WPI: 2000-611515/58.
 DR P-PSDB: AAB58845.
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -

Claim 1; Page 581-582; 1299pp; English.

Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neurotrophic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antibacterial; antiinflammatory; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and agonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 SO Sequence 917 BP; 228 A; 203 C; 270 G; 211 T; 5 other:

alignment_scores: Length: 96
 Quality: 170.00 Gaps: 2
 Ratio: 2.656
 Percent Similarity: 66.667 Percent Identity: 39.583

alignment_block:

US-09-327-750D-34 x AAF21748 ..

Align seg 1/1 to: AAF21748 from: 1 to: 917

```

29 GlnSerGluGluSerHisLsleuGluValGluAsnLysLysPr 45
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
336 CAGATGAGAGAGAGAGCGCCCTTGGAGAGAGTGAAGCCACGACC 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
45 oGlyGlyAsnValArgGlyValArgGlyValProAsnPhelLeut 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
386 TGCAGGAATGACAGGAGAGGCTGCGGACTTCCCTAAATTTTGCAT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 rPalalLeProAsnArgHisValAspHisSerGly.....GlyGlu 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
436 GGGCCATACCCATAGGAGATGATGATGATGATGATGATGATGAT 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
77 GluValLysArgPheValGlyValMetGluAlaLysArgHisSerly 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

486 GATATGGAATATTCATGGAGGAGATGAGAGAAATCAGAGAAACACTTAG 535
93 sGluGlnInMetArgProTyrThrArgPheArgThrProGluProAspA 110
536 GGAGCTGCGATTGAGGAATGCTCGGTATCTTATGGGGGAGCTCTCTA 585
110 snHisTyrAsp.....PheCysLeuIlePro 118
586 ATCACCATGACCATCATGATGAATTTGCTTATGCCTTATGCCT 623

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2001.DAT.AAF23528

seq_documentation_block:

ID AAF23528 standard; DNA; 700 BP.

XX AC AAF23528;

XX DT 22-MAR-2001 (first entry)

XX DE Mouse NADE DNA.

XX KW Neurotrophin receptor; p75-NTR; NGF-induced apoptosis;
XX KW neurogenetic disease; NF-kappaB; ds.

XX OS Mus sp.

XX PN WO200075278-A2.

XX PD 14-DEC-2000.

XX PF 07-JUN-2000; 2000WO-US15621.

XX PR 07-JUN-1999; 99US-0327750.

XX PA (UYCO) UNIV COLUMBIA NEW YORK.

XX PI Sato T;

XX DR WPI; 2001-061707/07.

XX PT New p75-neurotrophin receptor-associated cell death executor (NADE) and
XX PT the gene encoding NADE, useful for modulating the activity of p75NTR
XX PT and for detecting neurodegenerative diseases.

XX PS Claim 12; Fig 1; 134pp; English.

XX CC The present invention relates to a purified polypeptide capable of
XX CC binding neurotrophin receptor (p75-NTR). The invention is useful for
XX CC binding and modulating the activity of p75NTR. The peptide mediates
XX CC NGF-induced apoptosis, which plays an important role in neurogenetic
XX CC diseases. The peptide of the invention and p75NTR are useful for
XX CC inhibiting NF-kappaB activation in a cell or a subject, for inducing
XX CC caspase-2 and caspase-3 activity to cleave poly (ADP-ribose) polymerase
XX CC and fragment nuclear DNA in a cell by co-expression of (I) and p75-NTR.

XX SQ Sequence 700 BP; 177 A; 188 C; 203 G; 132 T; 0 other;

alignment_scores:

Quality: 144.00 Length: 129
Ratio: 2.000 Gaps: 3
Percent Similarity: 55.814 Percent Identity: 28.682

alignment_block:

US-09-327-750D-34 x AAF23528 ..

Align seg 1/1 to: AAF23528 from: 1 to: 700

18 LysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlucose 34
162 AAACAATAATCTCATCTGGCCATGTCACCCAGGAAACAGAGAGCTGGA 211
34 rHisHisLeuGluGluValGluAsnLysLysProGlyGly..... 47

212 GCACCCCTGCAGATGAGCAGGAGAACCCGCTGTGGGAGGAGGTGAGG 261
47
262 GCCACCAGCCTGTGCAACAAACAACAACAACAACAACAACAACAACAACCAC 311
48AsnValArgArgLysValArgLeuValProAsnPheLeuTr 62
312 AACCAACCACCAAGAGCCAGGCTCGCCGACTTGCCTTAACCTCCGATG 361
62 PALAileProAsnArgHisValAspHisSerGluGly.....GlyGluG 77
362 GGCATTCCCAACAGGAGATGAATGACGGTGGTGGGAGATGGAGATG 411
77 luValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
412 ATATGGAATCTTCATGAGGAGATGAGAGATCCGGAGAAAGCTTAGG 461
94 GluGlnInMetArgProTyrThrArgPheArgThrProGluProAspAs 110
462 GAGTACAGCTGAGAAATGCTACGCATCCTTATGGGGGAGCTGTCTAA 511
110 nHisTyrAsp.....PheCysLeuIlePro 118
512 CCACCAGCATCACCATCATGATGAATTTGCTTATGCCT 548

seq_name: /SDS2/gcgdata/geneseq/geneseq/NA2000.DAT.AAC01004

seq_documentation_block:

ID AAC01004 standard; cDNA; 485 BP.

XX AC AAC01004;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 1002.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GBST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR P-FSDB; AAG00998.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX XX Claim 1; SEQ ID 1002; 71pp + CD-ROM; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX CC well suited for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been
XX CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from


```
OS Homo sapiens.
XX CN1296969-A.
XX PD 30-MAY-2001.
XX PF 23-NOV-1999; 99CN-0124078.
XX PR 23-NOV-1999; 99CN-0124078.
XX PA (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX PI Mao Y, Xie Y;
XX WPI: 2001-483897/53.
DR P-PSDB; AAG66407.
XX Polypeptide-human X chromosome linked gene expression protein 14 and
XX polynucleotide for coding said polypeptide -
XX Claim 6; Page 24 (Disclosure); 31pp; Chinese.
XX The present sequence is the coding sequence for human X chromosome linked
XX gene expression protein 14. The protein and coding sequence are useful
XX for treating diseases e.g. cancer and HIV infection.
XX Sequence 767 BP; 224 A; 143 C; 215 G; 185 T; 0 other;

alignment_scores:
Quality: 111.00 Length: 132
Ratio: 1.461 Gaps: 6
Percent Similarity: 57.576 Percent Identity: 31.818

alignment_block:
US-09-327-750D-34 x AAH75810 ..
Align seg 1/1 to: AAH75810 from: 1 to: 767

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
101 ATGAGTCCAAAGAGGAGGAGCGTTAAACAATCTCATCGTGGAAATGT 150
15 .....GluLysAspLysLysLysLysLysGlyGlyLysAlaSerLysG 29
151 CAACGAGGAAATGATGAAAGAGATGAAAGAGGAGCAAGTGTGTAATAA. 199
29 InSerGluGluGluSerHisHisLysGluGluValGluAsnLysLysPro 45
200 ..GGGAGGCCCTTGGCCCTACCTTTCGAATGTTAGTGAATACGTGTCCT 247
46 GlyGlyAsnValArgArg....LysValArgArgLeuValProAsnPheLe 61
248 AGAGGAACCGTAGGCGGTTCGCGTTAGGAGCCCATCCTCGCATATAG 297
61 uTPAlaIleProAsnArgHisValAsp.....HisSerG 73
298 ATGGACATAATCATAGGCTTGGAGCCACAGCAAGGATGAGAGAGG 347
73 LuGlyGlyGluGluValGlyArgPheValGlyValMetGluAlaLys 89
348 AGAATATGGAAGATTTGGGAGGAGGTGACAGAGCTGATGGAAGCTG 397
90 ArgHisSerLysGluGlnGlnMetArgProTyrThrArgPheArgThrPr 106
398 AGG.....GAAAGACAGTTGAGTCATAGTCTGCGGCGAGTCAGCAC 438
106 oGluPro.....AspAsnHisTyrAspPheCysLeuIlePro 118
439 TGATCCCCCTCACCATGACCATCAGCATGATGTTGCTTATGCC 484

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT:AAI58581
```

```
seq_documentation_block:
ID AAI58581 standard; cDNA; 862 BP.
XX AC AAI58581;
XX DT 22-OCT-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 784.
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX anyotrophic lateral sclerosis; Shy-brager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; seq
XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US34263.
XX PR 21-JAN-2000; 2000US-0488725.
XX PR 25-APR-2000; 2000US-0552317.
XX PR 09-JUL-2000; 2000US-0598042.
XX PR 19-JUL-2000; 2000US-0620312.
XX PR 03-AUG-2000; 2000US-0653450.
XX PR 14-SEP-2000; 2000US-0662191.
XX PR 19-OCT-2000; 2000US-0693036.
XX PR 29-NOV-2000; 2000US-0727344.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
XX P-PSDB; AAM39425.
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX Claim 1; SEQ ID NO 784; 10078pp; English.
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-brager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.
XX Note: The sequence data for this patent did not form part of the printed
XX specification.
XX Sequence 862 BP; 235 A; 183 C; 238 G; 206 T; 0 other;

alignment_scores:
Quality: 111.00 Length: 132
Ratio: 1.461 Gaps: 6
Percent Similarity: 57.576 Percent Identity: 31.818

alignment_block:
```

US-09-327-750D-34 x AAF58581 ..

Align seg 1/1 to: AAF58581 from: 1 to: 862

```

1 MetalaserlyValysGlnValilleuAspleuthrVal..... 14
  ||| ||||| ..... ||| |||
  226 ATGAGCTCCAAAGAGACAGCAGCTTAAACATCATCTAGCGAAATGT 275
  15 ..... GlutylAspLysLysAsnLysLysGlyGlyLysAlaserlysg 29
  ||| ||||| ..... ||| |||
  276 CAACGAGAAATGATGATAAAGATGATAAAGAGCAAGTCTATATAA. 324
  29 InserlglugluginuserHshistleuglgluValgluAsnLysLysPro 45
  ||| ||||| ..... ||| |||
  325 ..GGGAGCCCTGGCCCTACCTTGAAATGTTAGTGAATCTGTGCT 372
  46 GlyLysAsnValArgArg...LysValArgArgLysValProAsnPhle 61
  ||| ||||| ..... ||| |||
  373 AGAGGAAACCGTAGCGGCTCCGCTAGCGACCCATCTCGATAG 422
  61 utrpAlaileProAsnArgHshValasp.....Hisserg 73
  ||| ||||| ..... ||| |||
  423 ATGCGACATATATGATAGCTTGAGAGCCACAGCAGAGTAGAGAGG 472
  73 InlglYlglYlgluValglYlArgPheValglYlGlnValMetGluAlaLys 89
  ||| ||||| ..... ||| |||
  473 AGAATATGAAAGATGGGAGAGAGTAGAGACAGCTGATGAAAGCTG 522
  90 ArgHshSerLysGluGlnMetArgProTyrThrArgPheArgThrPr 106
  ||| ||||| ..... ||| |||
  523 AGG.....GAAAGAGTTGAGTCATAGTTGGGAGGAGACACAC 563
  106 ogLupro.....AspAsnHshTyrAspPheCysLeuIlePro 118
  ||| ||||| ..... ||| |||
  564 TGTATCCCTCCATGACATGACATGACATGATGATTTGCTTATGCC 609
  seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA2001.DAT:AAF59611
  seq_documentation_block:
  ID AAF59611 standard; cDNA; 898 BP.
  AC AAF59611;
  XX
  XX
  DT 24-APR-2001 (first entry)
  XX
  XX
  DE Human cell cycle and proliferation protein CCYPR-22 cDNA, SEQ ID NO:76.
  XX
  XX
  KW Cell cycle and proliferation protein; CCYPR; human; agonist;
  KW antagonist; gene therapy; detection; gene therapy;
  KW transgenic animal disease model; immune disorder;
  KW developmental disorder; cell signalling disorder;
  KW cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
  KW arteriosclerosis; asthma; allergy; diabetes mellitus;
  KW menstrual cycle disorder; bacterial infection; ss.
  XX
  XX
  OS Homo sapiens.
  XX
  XX
  PN WO200107471-A2.
  XX
  PD 01-FEB-2001.
  XX
  XX
  PF 21-JUL-2000; 2000MO-US19948.
  XX
  XX
  PR 21-JUL-1999; 9905-0145075.
  PR 08-SEP-1999; 9905-0153129.
  PR 10-NOV-1999; 9905-0164647.
  XX
  XX
  PA (INCY-) INCYTE GENOMICS INC.
  XX
  PI Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O,
  PI Azimzai Y, Yang J, Lu DAM, Baughn MR, Patterson C, Shah P,
  XX
  DR WPI; 2001-112727/12.

```

DR P-PSDB; AAB60474.

Human cell cycle and proliferation proteins and polynucleotides are used to treat, diagnose and prevent immune, developmental and cell signalling disorders and cell proliferative disorders including cancer - Claim 5; Page 181-182; 205pp; English.

Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human cell cycle and proliferation proteins (CCYPR). AAB60453-AAB60506. CCYPR and agonists of CCYPR are used to treat diseases or conditions associated with decreased expression of functional CCYPR, while CCYPR antagonists are used to treat diseases or conditions associated with overexpression of functional CCYPR. Monoclonal or polyclonal antibodies to CCYPR may be used in enzyme-linked immunosorbent assays (ELISA) or radioimmunoassays to detect CCYPR. CCYPR itself may be used to detect compounds e.g., antibodies, oligonucleotides and proteins (receptors) that specifically bind to CCYPR, and in drug screening methods to identify compounds that modulate the activity of CCYPR. CCYPR nucleotides can be used to generate transgenic animal models of human disease, and can be used in gene therapy in target cells with genetic abnormalities with respect to the expression of CCYPR for the treatment or prevention of a disorder associated with CCYPR. Diseases which can be diagnosed, treated and prevented using CCYPR proteins, nucleic acids, agonists or antagonists include immune, developmental and cell signalling disorders, and cell proliferative disorders including cancer. Specific examples of these disorders include anaemia, epilepsy, arteriosclerosis, asthma, cancer, allergies, diabetes mellitus, disorders of the menstrual cycle and infections caused by bacteria.

Sequence 898 BP; 250 A; 186 C; 251 G; 211 T; 0 other;

alignment_scores:
 Quality: 111.00 Length: 132
 Ratio: 1.461 Gaps: 6
 Percent Similarity: 57.576 Percent Identity: 31.818

alignment_block:

US-09-327-750D-34 x AAF59611 ..

Align seg 1/1 to: AAF59611 from: 1 to: 898

```

1 MetalaserlyValysGlnValilleuAspleuthrVal..... 14
  ||| ||||| ..... ||| |||
  254 ATGAGCTCCAAAGAGACAGCAGCTTAAACATCTATCGTGAAATGT 303
  15 ..... GlutylAspLysLysAsnLysLysGlyGlyLysAlaserlysg 29
  ||| ||||| ..... ||| |||
  304 CAACGAGAAATGATGATAAAGATGATAAAGAGCAAGTCTATATAA. 352
  29 InserlglugluginuserHshistleuglgluValgluAsnLysLysPro 45
  ||| ||||| ..... ||| |||
  353 ..GGGAGCCCTGGCCCTACCTTGAAATGTTAGTGAATCTGTGCT 400
  46 GlyLysAsnValArgArg...LysValArgArgLysValProAsnPhle 61
  ||| ||||| ..... ||| |||
  401 AGAGGAAACCGTAGCGGCTCCGCTAGGAGCCCATCTGCGATAG 450
  451 ATGAGGACATATGATAGCTTGAGAGCCACAGCAGAGTAGAGAGG 500
  61 utrpAlaileProAsnArgHshValasp.....Hisserg 73
  ||| ||||| ..... ||| |||
  451 ATGAGGACATATGATAGCTTGAGAGCCACAGCAGAGTAGAGAGG 500
  73 InlglYlglYlgluValglYlArgPheValglYlGlnValMetGluAlaLys 89
  ||| ||||| ..... ||| |||
  501 AGAATATGAAAGATGGGAGAGAGTAGAGACAGCTGATGAAAGCTG 550
  90 ArgHshSerLysGluGlnMetArgProTyrThrArgPheArgThrPr 106
  ||| ||||| ..... ||| |||
  551 AGG.....GAAAGAGTTGAGTCATAGTGTGCGGAGCAGTACACAC 591
  106 ogLupro.....AspAsnHshTyrAspPheCysLeuIlePro 118

```


XX Sequence 662 BP; 177 A; 144 C; 192 G; 137 T; 12 other;
 SO Alignment scores:
 Quality: 103.00 Length: 132
 Ratio: 1.411 Gaps: 6
 Percent Similarity: 55.303 Percent Identity: 30.303

alignment_block:
 US-09-327-750D-34 x AAC03880 ..
 Align seg 1/1 to: AAC03880 from: 1 to: 662

```

1 MetAlaSerLysValLysGlnValLlleuAspLeuThrVal..... 14
207 ATGAGATCCAAAGAGAACGAGCGCTTAACAAATCTCATCTGGAATAATGT 256
15 .....GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysG 29
257 CAACCGAGAAATGATGAAAAAGATGAAAAGAGCAAGTGCCTAATMAA. 305
29 InSerGluGluGlnSerHisHisLeuGluGluValGluAsnLysLysPro 45
306 .....GGGAGGCCCTTGCCCTACCTTGTGATGTGTGTAATCTGTGCCT 353
46 GlyGlyAsnValArgArg..LysValArgArgLeuValProAsnPhel 61
354 AGAGCAAAATCTGAGCGCGTTCGCGTAGGAGCCCTCTCGAGATAG 403
61 uTPAlaIleProAsnArgHisValAsp.....HisSerG 73
404 ATGGGATATGATGATAGCTTAGGCTTGAGAACCAAGCAAGCATGANNAG 453
73 LuGlyGluGluValAlaLysArgPheValGlyGlnValMetGluAlaLys 89
454 AGAATATGGAAGAGATGGGAGAGTGAACAGACGTGATGAAAAAGCTG 503
90 ArgHisSerLysGluGlnMetArgProTyrThrArgPheArgThrPr 106
504 AGG.....GAAAGCAGTGTGATGCTGTGCGGCGCATGCAGCAC 544
106 ogLupro.....AspAsnHisTyrAspPheCysLeuIlePro 118
545 TGACCCCTCACCATGACATCATGATGATTTGGCNMNAATCC 590
seq_name: /SID2/gcdata/geneseq/geneseqn/NA2000.DAT:AAC03879
seq_documentation_block:
ID AAC03879 standard; CDNA: 698 BP.
XX AAC03879;
AC
XX
XX 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 3877.
XX
XX
XX Human: 5' EST: expressed sequence tag: secreted protein; CDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP103401-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-0200610.
PF
XX
XX 26-FEB-1999; 99US-0122487.
PR
XX (GEST ) GENSET.
PA
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
```

XX WPI; 2000-500381/45.
 DR P-PSDB; AAG03873.
 DR
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1; SEQ ID 3877; 71pp + CD-ROM; English.

XX The present sequence is one of a large number of 5' ESTs derived from
 CC cDNAs encoding secreted proteins. An ORF has been identified within the
 CC mRNA sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.

XX
 SO Sequence 698 BP; 187 A; 151 C; 210 G; 139 T; 11 other;

Alignment scores:
 Quality: 103.00 Length: 132
 Ratio: 1.411 Gaps: 6
 Percent Similarity: 55.303 Percent Identity: 30.303

alignment_block:
 US-09-327-750D-34 x AAC03879 ..
 Align seg 1/1 to: AAC03879 from: 1 to: 698

```

1 MetAlaSerLysValLysGlnValLlleuAspLeuThrVal..... 14
243 ATGAGATCCAAAGAGAACGAGCGCTTAACAAATCTCATCTGGAATAATGT 292
15 .....GluLysAspLysLysAsnLysLysGlyGlyLysAlaSerLysG 29
293 CAACCGAGAAATGATGAAAAAGATGAAAAGAGCAAGTGCCTAATMAA. 341
29 InSerGluGluGlnSerHisHisLeuGluGluValGluAsnLysLysPro 45
342 .....GGGAGGCCCTTGCCCTACCTTGTGATGTGTGTAATCTGTGCCT 389
46 GlyGlyAsnValArgArg..LysValArgArgLeuValProAsnPhel 61
390 AGAGCAAAATCTGAGCGCGTTCGCGTAGGAGCCCTCTCGAGATAG 439
61 uTPAlaIleProAsnArgHisValAsp.....HisSerG 73
440 ATGGATATGATGATAGCTTAGGCTTGAGAACCAAGCAAGCATGANNAG 489
73 LuGlyGluGluValAlaLysArgPheValGlyGlnValMetGluAlaLys 89
490 AGAATATGGAAGAGATGGGAGAGTGAACAGACGTGATGAAAAAGCTG 539
90 ArgHisSerLysGluGlnMetArgProTyrThrArgPheArgThrPr 106
540 AGG.....GAAAGCAGTGTGATGCTGTGCGGCGCATGCAGCAC 580
106 ogLupro.....AspAsnHisTyrAspPheCysLeuIlePro 118
581 TGACCCCTCACCATGACATCATGATGATTTGGCNMNAATCC 626
seq_name: /SID2/gcdata/geneseq/geneseqn/NA2001.DAT:AA160367
seq_documentation_block:
```


Align seg 1/1 to: AAI60367 from: 1 to: 858

1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal.....14
215 ATGGAGTCCAAAGAGGAGCGTTAAACAATCTCATCGTGGAAATGT 264
15GluLysAspLysLysLysGlyLysAlaSerLysG 29
265 CAACAGGAAATGATCAAAAGATGAAAGAGGAGCAAGTTGCTAATAA. 313
29 InSerGluGluSerHisHisLysGluValGluAsnLysLysPro 45
314 ..GGGAGGCCCTTGGCCCTACCTTTGAATGTTAGTGAATACTGTGTCCT 361
46 GlyGlyAsnValArgArg...LysValArgArgLeuValProAsnPheLe 61
362 AGAGGAAACCGTAGGCGGTTCCGCTTAGGCAGCCCATCTGCAGTATAG 411
61 uTrpAlaLysProAsnArgHisValAsp.....HisSerG 73
412 ATGGACATATGATGCTAGGCTTGGAGAGCCACAGGCAAGATGAGAGG 461
73 LuGlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAlaLys 89
462 AGAATATGGAAGGATGGGAGGAGGTGAGACAGCTGATGGAAGAGCTG 511
90 ArgHisSerLysGluGlnMetArgProTyrThrArgPheArgThrPr 106
512 AGG.....GAAAGCAGTTGAGTCATAGTCTGGGGCAGTCAGCAC 552
106 oGluPro.....AspAsnHisTyrAspPheCys 115
553 TGATCCCTCCATCAGCATCAGATGAGTTTGC 589

AAI60367 standard; cDNA; 858 BP.
AAI60367;
22-OCT-2001 (first entry)
Human polynucleotide SEQ ID NO 4356.
Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia; ss.
Homo sapiens.
WO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US34263.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
P-PSDB; AAM41211.
Novel nucleic acids and polypeptides, useful for treating disorders
such as central nervous system injuries -
Claim 1; SEQ ID NO 4356; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and
the encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
Note: The sequence data for this patent did not form part of the printed
specification.
Sequence 858 BP; 236 A; 175 C; 241 G; 206 T; 0 other;

alignment_scores:
Quality: 99.00 Length: 129
Ratio: 1.356 Gaps: 6
Percent Similarity: 56.589 Percent Identity: 31.008
alignment_block:
US-09-327-750d-34 x AAI60367

OM of: US-09-327-750D-34 to: GenEmbl.* out_format : pfs

Date: Mar 11, 2002 3:34 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODE=frame-p2n.model -DEV=xlh
-q/-cgn2_170sp100sp100/US09327750/runat_11032002_101154_20319/app_query.fasta_1.1472
-BB-GenEmbl -qfmt=fastap -suffix=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEX=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELXT=7.000 -START=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09327750 -CGN1_1.8673
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-327-750D-34

Query length: 118

Database: GenEmbl.*

Database sequences: 1472140

Database length: 341344837

Search time (sec): 4557.230000

score_list:

Sequence	Strd	Orig	ZScore	Escore	Len	Documentation
gb.pr:AK000959	+	372.00	661.44	1.4e-28	1229	AX000959 Homo sapiens CDNA FLJ
gb.pr:AX100231	+	372.00	660.55	1.4e-28	1364	AX100231 Sequence 32 from Pat
gb.pr:HS635619	+	372.00	627.19	1.2e-26	69648	AX035494 Human DNA sequence f
gb.pr:HUMOGC	+	171.50	302.23	1.5e-08	891	M3188 Human unknown protein f
gb.pr:AF187064	+	171.50	302.23	1.5e-08	891	AF187064 Homo sapiens p75NTR-a
gb.pr:BC003190	+	170.00	300.51	1.8e-08	793	BC003190 Homo sapiens p75NTR-a
gb.ro:AF187085	+	148.00	264.39	1.9e-06	519	AF187085 Rattus norvegicus p75N
gb.sts:G72708	+	145.00	260.75	3.0e-06	421	G72708 MARC 4953-4954:199319031
gb.ro:AF187066	+	145.00	256.44	5.2e-06	700	AF187066 Mus musculus p75NTR-a
gb.ro:AF097440	+	145.00	254.75	6.5e-06	854	AF097440 Mus musculus brain exp
gb.pr:HSV351F8	+	138.00	208.36	0.0025	45678	G35294 human STS SHGC-37409, se
gb.sts:G35294	+	118.50	211.86	0.0016	477	AF097439 Mus musculus brain exp
gb.ro:AF097439	+	116.50	204.02	0.0044	785	AF183416 Homo sapiens ovarian g
gb.pr:AF183416	+	112.50	198.50	0.0088	642	AF220189 Homo sapiens uncharact
gb.pr:AF220189	+	112.50	196.35	0.0117	828	AF220189 Homo sapiens uncharact
gb.pr:HS198P4	+	112.50	164.41	0.6398	35714	AF1008708 Human DNA sequence f
gb.pr:AF251053	+	111.00	194.55	0.0147	744	AF251053 Homo sapiens X-linked
gb.pr:AX078272	+	111.00	192.95	0.0180	898	AX078272 Sequence 76 from Paten
gb.pr:HSV87088	+	111.00	162.82	0.8586	31321	AF20233 Human DNA sequence f
gb.pr:AL133348	+	111.00	160.62	1.14	40584	AL133348 Human DNA sequence f
gb.pr:HSU080B1	+	111.00	160.53	1.15	41029	AF022169 Homo sapiens chromos
gb.pr:AF237783	+	109.50	191.32	0.0222	791	AF237783 Homo sapiens brain exp
gb.pr:AF097438	+	105.00	182.74	0.0668	835	AF097438 Mus musculus brain exp
gb.pr:AF051347	+	105.00	182.71	0.0670	838	AF051347 Mus musculus REX-3 mRN
gb.pr:AF097437	+	105.00	174.26	0.1981	2269	AF097437 Mus musculus Bex1 prc
gb.pr:AF0714B7	+	102.50	137.78	21.32	1225	AF0714B7 Streptococcus pneumon
gb.pr:AF071813	+	96.00	163.24	0.8138	1822	AF071813 Homo sapiens chromo
gb.pr:AC023561	+	95.50	120.50	195.44	169808	AC023561 Homo sapiens chromo
gb.pr:AC074348	+	95.00	120.24	202.20	157511	AC074348 Homo sapiens clone
gb.pr:AC048344	+	95.00	118.73	232.92	179380	AC048344 Homo sapiens chromo
gb.pr:AC023341	+	94.50	119.12	245.66	169365	AC023341 Homo sapiens chromo
gb.pr:AC023341	+	94.00	130.14	56.75	39602	AC000373 Human Cosmid g1862d2
gb.pr:HSAC000373	+	94.00	129.38	62.62	43349	AC000367 Human Cosmid g1862x0
gb.pr:AC012232	+	94.00	117.57	284.63	174337	AC012232 Homo sapiens clone
gb.pr:HS735P9	+	92.50	118.28	255.95	116361	AL032633 Human DNA sequence
gb.pr:AL355809	+	92.50	118.01	269.13	120340	AL355809 Human DNA sequence
gb.pr:AF195953	+	92.50	113.34	489.70	208618	AF195953 Homo sapiens membra
gb.in:CSA7431	+	92.00	151.74	3.56	2029	AJ007431 Cupiennius salei mRNA
gb.pr:AL512842	+	92.00	114.96	397.64	154889	AL512842 Homo sapiens chromo
gb.pr:AC091457	+	92.00	112.30	559.73	212080	AC091457 Mus musculus clone

gb.btg:AL139096 + 91.50 115.47 372.68 131198 ! AL139096 Homo sapiens chr
gb.pr:U72787 + 91.00 131.18 49.69 18510 ! U72787 Homo sapiens cosmid
gb.pr:AL359643 + 91.00 112.53 543.53 166863 ! AL359643 Human DNA sequen
gb.pr:HS6278N14 - 91.00 111.78 597.92 182150 ! AL109654 Human DNA sequen
gb.btg:B1011H02 + 90.50 119.60 219.49 65190 ! AL442109 Oryza sativa chro

seq_name: gb.pr:AK000959

seq_documentation_block: 1229 bp mRNA PRI 22-FEB-2000
LOCUS AK000959
DEFINITION Homo sapiens CDNA FLJ10097 fis, clone HEMBA1002458, weakly similar
to OVARIAN GRANULOSA CELL 13.0 KD PROTEIN HGR74.

ACCESSION AK000959

VERSION GI:7021945

KEYWORDS oligo tapping; fis (full insert sequence).

SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to

MRNA, clone_lib:HEMBA1 clone:HEMBA1002458.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,K., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Oshima,A.
NEDO human cDNA sequencing project

Unpublished (2000)

2 (bases 1 to 1229)

Isogai,T. and Otsuki,T.

Direct Submission

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

International Trade and Industry of Japan; cDNA full insert

sequencing; Research Association for Biotechnology; cDNA library

construction, 5'- & 3'-end one pass sequencing and clone selection;

Helix Research Institute (supported by Japan Key Technology Center

etc.) and Department of Virology, Institute of Medical Science,

University of Tokyo.

FEATURES

source

1. .1229

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HEMBA1002458"

/dev_stage="embryo, 10 weeks"

/tissue_type="whole embryo, mainly head"

/note="cloning vector: pME18SFL3"

191. .553

/note="unnamed protein product"

/codon_start=1

/protein_id="BAA91443.1"

/db_xref="GI:7021946"

/translation="MESKEELAAANLNGENAAQENGEQAQPTONEEESRHGGEGGQ

KPGNTRRRGRVRLVPLNFRWALPNRHNHNEARDVERFVGOMMEIKRKTREQQMRHY

MRPQPEPDNHYDECLIP"

BASE COUNT 351 a 243 c 296 g 339 t

ORIGIN

alignment_scores:

Quality: 372.00 Length: 120

Ratio: 3.758 Gaps: 2

Percent Similarity: 82.500 Percent Identity: 60.000

alignment_block:

US-09-327-750D-34 x AK000959

Align seg 1/1 to: AK000959 from: 1 to: 1229

```

17  p1ys1ysa1n1ys1ys1g1y1... 1ys11as1er1ys1n1ser1g1u1g 33
    .....
300  CCAGACAGAGAAACGAGAGAGGGGAGCGAGGCCCCACCGACGACGATGAGAGAG 349
    .....
33  1use1his1ist1euc1u1u1val1g1u1a1n1ys1n1ys1p1ro1g1y1a1n1val 49
    .....
350  AATCCGCCCATTTGGAGGAGGGGTGAGAGCCGACGAGCGCTGAGAGAAATATC 399
    .....
50  ATGATG... 1ys1val1ar1g1le1u1val1p1ro1as1phe1u1tr1p1al1le1p1r 65
    .....
400  AGGCGGGGGCGAGTTAGCGGACTTGTCCCAATTTTGCATGGGCCCATACC 449
    .....
65  oas1nar1h1is1a1s1p1h1is1ser1g1u1g1y1g1u1g1u1val1g1ar1g1p1he1y 82
    .....
450  TAATAGCATATTGAGACCAATGAAGCAGAGATGATGAGAAAGTTTG 499
    .....
82  a11g1n1val1met1g1u1a1l1ys1ar1g1h1is1er1ys1g1u1g1n1et1ar1g 98
    .....
500  TAGGCGACATGATGGAATCAAGAGAAAGACTAGGAGACGAGAGAGAG 549
    .....
99  p1ro1y1tr1h1ar1g1p1he1r1h1r1p1ro1u1p1ro1k1p1s1p1h1is1ty1r1as1p1p1he1y 115
    .....
550  CACATATATGGCTTCACAACTCTCAACCTCACACACACATATGACTTTGG 599
    .....
115  sleu1le1p1ro 118
    .....
600  CCTCATACCT 609
    .....
seq.name: gb_pr:HS635G19

seq.docmentation block:
LOCUS      HS635G19      69648 bp                PRI      23-NOV-1999
DEFINITION Human DNA sequence from clone 635G19 on chromosome Xg22.1-22.3
            Contains a LMRI (laminin Receptor 1 (67KD) (RPSA, 40S Ribosomal
            Protein S $\alpha$ , P40)) pseudogene and part of a novel protein. Contains
            ESTs and GSSs, complete sequence.
ACCESSION  AL035494
VERSION    AL035494.8  GI:4775622
KEYWORDS   HTG: 40S Ribosomal Protein P40; 40S Ribosomal Protein S $\alpha$ ; CPG
            Island; Laminin Receptor 1; LMRI; RPSA.

SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69648)
Direct Submission
Submitted (02-JUN-1999) Sanger Centre, Hinxton, Cambridgeshire,
UK
SBUID16, UK. E-mail enquiries: hnmquery@sanger.ac.uk
Request: clonequests@sanger.ac.uk
On May 11, 1999 this sequence version replaced gi:4678749.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations.
Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unseq'
feature key.
This sequence was generated from part of bacterial clone configs of
human chromosome X, constructed by the Sanger Centre Chromosome X
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChX
635G19 is from the library PC14 constructed at the Roswell Park
Cancer Institute by the group of Pieter de Jong. For further

```


Authors	Title
Kapp, G., Freudenstein, O., Zimmer, M., and Scheit, K.H.	Characterization of three abundant mRNAs from human granulosa cells

```

113 p.....PhcysLeu11Pro 118
      |          |||||||:::|

```

```
Location/Qualifiers
1
/organism="Homo sapiens"
/db_xref="LocusID:37018"
/db_xref="taxon:9606"
/clone="MGC:802 IMAGE:3357965"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NH_MGC_16"
/lab_host="PH10B-R"
/note="Vector"

```

```
Location/Qualifiers
1
/organism="Homo sapiens"
/db_xref="LocusID:37018"
/db_xref="taxon:9606"
/clone="MGC:802 IMAGE:3357965"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NH_MGC_16"
/lab_host="PH10B-R"
/note="Vector"

```



```

|||||
162 AAAAAAAAAATCATCATGCGCATATGTCACCGAGAAAGAGATGGA 211
34 THSHISLeuGluGluValGluAsnLysLysProGlyGly..... 47
212 GCACCCCTCTGAGATGAGAGAGAGACCCCTCTGTGGAGAGGTGAG 261
47 ..... 47
262 GCCACGAGCGCTGTCGAAACACACACACACACACACACATTAACAC 311
48 .....AsnValArgArgLysValArgArgLeuValProAsnPhleuTr 62
312 AACCAACCGAGAGCGCGAGCGCTGCGACCTTCCCTTAACCTCCAGT 361
62 PalaileproAsnArgHisValAspHisSerGlyGly.....GlyGluG 77
362 GGCCTTCCTCCACACGACGATGATGACGGTGGTGAGATGAGATG 411
77 LuValGlyArgPheValGlyValMetGluAlaLysArgHisSerLys 93
412 ATATGGAATGTCATGAGAGATGAGAGATGCGGAGAAAGCTTAGG 461
94 GluGluGlnMetArgProTyrThrArgPheArgThrProGluProAsp 110
462 GAGCTACGAGCTGAGAAATGCTACGACATCTTATGGGAGCTGTCTAA 511
110 nHISTyrASP.....PheCysLeuIlePro 118
512 CCACCGATCACCATGATGATTCCTTATGCTTATGCT 548

seq_name: gb_to:AF097440

seq_documentation_block:
LOCUS AF097440 854 bp mRNA ROD 13-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 3 (Bex3) mRNA,
complete cds.
ACCESSION AF097440
VERSION AF097440.1 GI:4580593
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 854)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Bex1, a gene with increased expression in parthenogenetic embryos,
is a member of a novel gene family on the mouse X chromosome
JOURNAL Hum. Mol. Genet. 8 (4), 611-619 (1999)
MEDLINE 99172070
REMARK Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 854)
AUTHORS Brown,A.L. and Kay,G.F.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
FEATURES
Source
1..854
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/db_xref="dbEST:AA272375"
/chromosome="X"
/map="near Pip"
/tissue_type="pooled organs"
1..854
/gene="Bex3"
/feature="Bex3"
172..546
/gene="Bex3"
/codon_start=1
/product="brain expressed X-linked protein 3"
/protein_id="AAD24431.1"

```

```

/dB_xref="GI:4580594"
/translation="MAVHVEEENEOPFQNGEDRPVGGEGHOPANNNNNNNHN
HNHRRGQARRLAPNFRWATPNRMDGSLGGDDMEFMEERIRKRLRELQALNC
LRIIMGELSHNHHDHDFCLMP"
BASE COUNT 237 a 212 c 228 g 177 t
ORIGIN

Alignment Scores:
Quality: 145.00 Length: 129
Ratio: 2.014 Gaps: 3
Percent Similarity: 55.814 Percent Identity: 28.682

alignment_block:
US-09-327-750D-34 x AF097440 ..
Align seg 1/1 to: AF097440 from: 1 to: 854

18 LysLysAsnLysLysGlyGlyLysAlaSerLysGlnSerGluGluGlu 34
|||||
157 AAAAAAAAAATCATCATGCGCATATGTCACCGAGAAAGAGATGGA 206
34 THSHISLeuGluGluValGluAsnLysLysProGlyGly..... 47
207 GCACCCCTCTGAGATGAGAGAGAGACCCCTCTGTGGAGAGGTGAG 256
47 ..... 47
257 GCCACGAGCGCTGTCGAAACACACACACACACACACATTAACAC 306
48 .....AsnValArgArgLysValArgArgLeuValProAsnPhleuTr 62
307 AACCAACCGAGAGCGCGAGCGCTGCGACCTTCCCTTAACCTCCAGT 356
62 PalaileproAsnArgHisValAspHisSerGlyGly.....GlyGluG 77
357 GGCCTTCCTCCACACGACGATGATGACGGTGGTGAGATGAGATG 406
77 LuValGlyArgPheValGlyValMetGluAlaLysArgHisSerLys 93
407 ATATGGAATGTCATGAGAGATGAGAGATGCGGAGAAAGCTTAGG 456
94 GluGluGlnMetArgProTyrThrArgPheArgThrProGluProAsp 110
|||||
457 GAGCTACGAGCTGAGAAATGCTACGACATCTTATGGGAGCTGTCTAA 506
110 nHISTyrASP.....PheCysLeuIlePro 118
|||||
507 CCACCGATCACCATGATGATTCCTTATGCTTATGCT 543

seq_name: gb_pr:HSV351F8

seq_documentation_block:
LOCUS HSV351F8 45678 bp DNA PRI 23-NOV-1999
DEFINITION Human DNA sequence from cosmid V351F8, between markers DXS366 and
DXS87 on chromosome X contains ESTs.
ACCESSION Z70719
VERSION Z70719.1 GI:1261915
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 45678)
AUTHORS Whiteley,M.
TITLE Direct Submission
JOURNAL Submitted (09-APR-1995) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1RQ, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT IMPORTANT: This sequence is the entire insert of clone V351F8. The
true left end of clone V351F8 is at 1 in this sequence. The true
right end of clone V351F8 is at 45678.
V351F8 is from the human chromosome X-specific cosmid library.
Location/Qualifiers
FEATURES

```

```

1. 45678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="X"
/clone="GHC-351F8"
/clone_lib="SCCV"
1..70
repeat_region
/partial
/note="Alu repeat: matches 80. .1 of consensus"
642. 930
repeat_region
/note="Alu repeat: matches 1. .308 of consensus"
1406. 5179
repeat_region
/note="L1 element fragment"
5243. 5276
repeat_region
/note="17 copies of 2 mer 82 % conserved"
5279. 5567
repeat_region
/partial
/note="Alu repeat: matches 304. .1 of consensus"
9975. 10369
repeat_region
/note="MSTA element fragment"
10399. 12017
repeat_region
/note="MSTAR element fragment"
11318. 11978
repeat_region
/note="THE1B element fragment"
11715. 11846
repeat_region
/note="MLTIR element fragment"
12074. 12170
repeat_region
/note="MSTA element fragment"
12227. 12486
repeat_region
/note="MSTA element fragment"
12437. 12502
repeat_region
/note="MSTC element fragment"
12756. 13059
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
14478. 14650
repeat_region
/note="THE1B element fragment"
14491. 14536
repeat_region
/note="MSTA element fragment"
14715. 14831
repeat_region
/note="THE1B element fragment"
15099. 15533
repeat_region
/note="L1 element fragment"
15534. 15626
repeat_region
/note="31 copies of 3 mer 85 % conserved"
15566. 15619
repeat_region
/note="3 copies of 18 mer 98 % conserved"
16352. 16436
repeat_region
/note="L1 element fragment"
16728. 16763
repeat_region
/note="9 copies of 4 mer 94 % conserved"
16764. 17054
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
19420. 19449
repeat_region
/note="15 copies of 2 mer 87 % conserved"
complement(19537..20013)
misc_feature
/note="match: 3' EST N51315 clone 283089"
20071. 20114
repeat_region
/note="22 copies of 2 mer 98 % conserved"
20073. 20112
repeat_region
/note="10 copies of 4 mer 100 % conserved"
20073. 20126
repeat_region
/note="3 copies of 18 mer 87 % conserved"
23088. 23300
repeat_region
/note="L1 element fragment"
23311. 23490
repeat_region
/note="MLTIR element fragment"
23318. 23489
repeat_region
/note="MLTIR element fragment"
23362. 23491
repeat_region
/note="MLTIR element fragment"
23507. 24407
repeat_region
/note="L1 element fragment"
24376. 24462
repeat_region
/note="MSTC element fragment"
24379. 24491
repeat_region
/note="MSTA element fragment"
24688. 24738
repeat_region
/note="MSTA element fragment"
24810. 25016
repeat_region
/note="L1 element fragment"
25254. 25421
repeat_region
/note="L1 element fragment"
25439. 25696
repeat_region
/note="MER25 element fragment"
26284. 26343
repeat_region
/note="MLTID element fragment"
26428. 26491
repeat_region
/note="MLTID element fragment"
26465. 26518
repeat_region
/note="MLTIE element fragment"
27446. 27667
repeat_region
/note="L1 element fragment"
28954. 29247
repeat_region
/partial
/note="Alu repeat: matches 308. .1 of consensus"
30160. 30325
repeat_region
/note="2 copies of 83 mer 98 % conserved"
32588. 32875
misc_feature
/note="match: 5' EST H68599 clone 239077"
complement(32825..33230)
misc_feature
/note="match: 3' EST H68239 clone 289077"
35551. 35802
repeat_region
/partial
/note="Alu repeat: matches 1. .260 of consensus"
35804. 35839
repeat_region
/note="18 copies of 2 mer 83 % conserved"
38625. 38705
repeat_region
/note="MLTIR element fragment"
38747. 38839
repeat_region
/note="MLTIR element fragment"
38949. 39032
repeat_region
/note="MLTIE element fragment"
38965. 39076
repeat_region
/note="MLTID element fragment"
BASE COUNT 14260 a 9135 c 9327 g 12956 t
ORIGIN
alignment_scores:
Quality: 138.00 Length: 127
Ratio: 2.000 Gaps: 3
Percent Similarity: 54.331 Percent Identity: 29.921
alignment_block:
US-09-327-750D-34 x HSV351F8
Align seg 1/1 to: HSV351F8 from: 1 to: 45678
12 LeuThrvAluGluLysAspLysLysLysLysLysGlyGlyLysAlaSerly 28
19408 TTATCTTCTAGGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGTCAACAT 19457
28 sGInser.....Glug 32
19458 GGAATATGTCCTCCCAAGAAAAACAAGATGTGTGAGAGAGCCCAAGTCAGA 19507
32 luGluSerHisLysLeuGluGluValGluAsnLysLysProGlyGlyAsn 48
19508 ATGAAGCCCGCTTTAGGAGGTGTGTGAATACCAAGGAGCTGTGAGGAAT 19557
49 ValArgArgLysValArgArgLeuValProAsnPheLeuTirPalalePr 65
19558 GTTAAAGGGGTTTGGGCTCCACCTGCCCCGGTGTGGAGAGGATGTGCC 19607

```

```

primer_bind      complement(130..149)      130 t
BASE COUNT      153 a      115 c      79 g
ORIGIN

alignment_scores:
    Quality:      118.50      Length:      83
    Ratio:        2.370      Gaps:      2
    Percent Similarity: 60.241      Percent Identity: 33.735

alignment_block:
005-09-32/-750D-34 x G35294/rev ..

Align seq 1/1 to reverse of: G35294 from: 1 to: 477

43 LysAspProGluGlyAsnValArgValArgValArgValProAs 59
   ::::::::::::::::::::::::::::
474 CAGAGAGCTGGAGGAAATGTTAAAGGGGTGGGCTCCACCTGCCCGG 425
59 nPhelLeuTrpAlaLeuProAsnArgHisValAspHisSerGlu..... 73
   ::::::::::::::::::::::::::::
424 TTTTGGAGAGAGATGTGCCCATAGCGCTTGCATTAACATGATGATAG 375
74 .....GlyGlyGluGluValGlyArgPheValGlyGlnValMetGluAla 88
   ::::::::::::::::::::
374 ATGAGAGATGGAGATGATATGGAGCGTTCTGAGAGAGATGAGAGAGCTA 325
89 LysArgHisSerLysGluGlnMetArgProTyrThrArgPheArgTyrH 105
   ::::::::::::::::::::
324 AGGAGGAAATTAAGGAACTTCAGTTGAGGTACAGCTCGCATCTTAT 275
105 rProGluPro.....AspAsnHisTyrAspPheCysLeuLeuPro 118
   ::::::::::::::::::::
274 AGGGAGCCCTCTCACCATGATCATCATGATGAGTTTGCTTATGCT 226

seq_name: gb_ro:AF097439

seq_documentation_block:
LOCUS      AF097439      785 bp      mRNA      ROD      13-APR-1999
DEFINITION Mus musculus brain expressed X-linked protein 2 (Bex2) mRNA.
complete cds.
VERSION      AF097439
KEYWORDS      AF097439.1 GI:4580591
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 785)
Brown,A.L. and Kay,G.F.
Bex2, a gene with increased expression in parthenogenetic embryos
is a member of a novel gene family on the mouse X chromosome
Hum. Mol. Genet. 8 (4), 611-619 (1999)
99172070
Erratum: [[published erratum appears in Hum Mol Genet 1999
May;8(5):943]]
2 (bases 1 to 785)
Brown,A.L. and Kay,G.F.
Direct Submission
Submitted (08-OCT-1998) Cancer Unit, Queensland Institute of
Medical Research, Herston Rd, Brisbane, Qld 4029, Australia
Location/Qualifiers
1..785
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="dbEST:W48832"
/chromosome="X"
/map="near Plp"
/dev_stage="embryo; 15.5 dpc"
1..785
/gene="Bex2"
139..528

```

```
/gene="Bex2"
/codon_start=1
/product="brain expressed X-linked protein 2"
/protein_id="AAD24430.1"
/db_xref="GI:4580592"
/translation="MESKEVEQGVKNLNMENDHOEKEEKPQDASKRDPVALPFEA
GDYVPGGRRERVRQPIVHYRWDLHMRVGEPOGRMEENVORFGDDVRQLMEKLE
RQLSHSLRAVSTDPHPHDDHDEFCLMP"
BASE COUNT      212 a 163 c 235 g 175 t
ORIGIN

alignment_scores:
  Quality: 116.50      Length: 130
  Ratio: 1.553         Gaps: 5
  Percent Similarity: 57.692      Percent Identity: 30.000

alignment_block:
  US-09-327-750D-34 x AF097439 ..
  Align seg 1/1 to: AF097439 from: 1 to: 785
    1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrVal..... 14
    139 ATGGAGTCCAAATGGCAACAGCGGTGAATAATCTCAACATGGAGAATGA 188
    15 .....GluLysAspLysLysLysLysGlyGlyLysAlaSerLysG 29
    189 CCATCAGGAAGAGGAGAAAGAAAGAAAGCCACAGGATGCTAGCAAAA 238
    29 InSerGluGluGluSerHisLysLysLysLysLysLysPro 45
    239 GGGATCGATGTGGCCCTGCTTCGAAGCTGGAGACTACTACGTGCT 288
    45 ...GlyGluValValArgLysValArgLysValArgLysValProAsnPheLe 61
    289 AGAGGAGGTGCGAGGGGTTCGGGGTTCGGAGCCCATCGTGCATACAG 338
    61 uTPAlaIleProAsnArgHisValAspHisSerGluGly.....G 75
    339 ATGGACCTGATGATAGG...GTTGGGGAGCCCGAGGAGGATGAGAG 385
    75 LyGluGluValValArgPheValGlyGlnValMetLysLysHis 91
    386 AGGAGACGTACAGAGGTGGGGATGATGATGAGACGCTCATGGAGAG 435
    92 SerLysGluGlnMetArgProTyrThrArgPheArgThrProGluPr 108
    436 CTGAGGGAAGGACGAGTGCACACAGCCTCGGGCGGTAGCACTGACCC 485
    108 o.....AspAsnHisTyrAspPheCysLeuIlePro 118
    486 GCCTCATGACCACTCATGATGATGTTTCCCTTATGCC 525

seq_name: gb_pr:AF183416
seq_documentation_block:
  LOCUS      AF183416      642 bp      mRNA      PRI      02-SEP-2000
  DEFINITION Homo sapiens ovarian granulosa cell 13.0 kDa protein hGR74 homolog
             mRNA, complete cds.
  ACCESSION  AF183416
  VERSION    AF183416.1 GI:9963770
  KEYWORDS   Unpublished
  SOURCE     human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 642)
  AUTHORS    Peng,Y., Qian,b., Tu,Y., Xu,S., Han,Z., Fu,G. and Chen,Z.
  TITLE      A novel gene expressed in human adrenal gland
  JOURNAL    Unpublished
  REFERENCE 2 (bases 1 to 642)
             Peng,Y., GU,W., Huang,C., Xu,S., Han,Z., Fu,G. and Chen,Z.

Direct Submission
Submitted (03-SEP-1999) Chinese National Human Genome Center at
Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
Shanghai 201203, P. R. China
FEATURES
  source
    1..642
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /tissue_type="adrenal gland"
    3..380
    /note="HGR74-h protein"
    /codon_start=1
    /evidence=not_experimental
    /product="ovarian granulosa cell 13.0 kDa protein hGR74
    homolog"
    /protein_id="AAG09685.1"
    /db_xref="GI:9963771"
    /translation="MESKEKRAVNSLSMENANOENKEEQVANKGEPLALPDAGEY
    C
    VPRGNRRFRVRQPIQYRWDMHRLGEPOARMRENMERIGEEVRLMEKLEKQLS
    HSLRAVSTDPHPHDDHDEFCLMP"
  CDS
    204 a 118 c 157 g 163 t
    BASE COUNT
    ORIGIN

alignment_scores:
  Quality: 112.50      Length: 128
  Ratio: 1.424         Gaps: 5
  Percent Similarity: 61.719      Percent Identity: 28.905

alignment_block:
  US-09-327-750D-34 x AF183416 ..
  Align seg 1/1 to: AF183416 from: 1 to: 642
    1 MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLys.. 16
    3 ATGGAGTCCAAAGAGAAAGAGGAGCAAGTGTCTAATAAAGGGGAGCCCT 52
    17 AspLysLysLysLysLysGlyGlyLysAlaSerLysLysGlnSerGluGlu 33
    53 CAACCAAGAAATGAAGAAAGAGGAGCAAGTGTCTAATAAAGGGGAGCCCT 102
    33 LuSerHisHisLeuGluGluValGluAsnLysLysProGlyGlyAsnVal 49
    103 TGGCCCTCCCTTGGATGCTGCTGAATACTGTGTCTAGAGGAAATCGT 152
    50 ArgArg...LysValArgArgLeuValProAsnPheLeuTyrAlaIlePr 65
    153 AGCGGCTTCCCGGTAGCGCCCTCTCTGCGATATAGATGGATATGAT 202
    65 oAsnArgHisValAsp.....HisSerGluGlyGlyGluG 77
    203 GCATAGGCTTGGAGAACCAACAGGCAAGGATGAGAGAGAGAAATGAAA 252
    77 LuValGlyArgPheValGlyGlnValMetGluAlaLysArgHisSerLys 93
    253 GGATGGGAGGAGGTGACAGAGCTGATGGAAAAGCTGAGG..... 293
    94 GluGlnGlnMetArgProTyrThrArgPheArgThrProGluPro..... 108
    294 GAAAGCAGTTGATCATAGTCTGCGGCGACGTCAGCAGTACCCCTCA 343
    109 .....AspAsnHisTyrAspPheCysLeuIlePro 118
    344 CATGACCATCATGATGATGTTTGCCTTATGCC 377

seq_name: gb_pr:AF220189
seq_documentation_block:
  LOCUS      AF220189      828 bp      mRNA      PRI      04-MAY-2000
  DEFINITION Homo sapiens uncharacterized hypothalamus protein HBEX2 mRNA,
             complete cds.
  ACCESSION  AF220189
```

508 CCATGACCATCATGATGAGTTTGCTTATGCC 541

```
VERSION      AF220189.1  GI:7689028
KEYWORDS
SOURCE
ORGANISM      human.
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS      Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE        A novel gene expressed in human hypothalamus
JOURNAL      Unpublished
REFERENCE
AUTHORS      Xiao,H., Song,H., Gao,G., Ren,S., Chen,Z. and Han,Z.
TITLE        Direct Submission
JOURNAL      Submitted (30-DEC-1999) Chinese National Human Genome Center at
               Shanghai, 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong,
               Shanghai 201203, P. R. China
               Location/Qualifiers
FEATURES
   source
       1..828
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /tissue_type="hypothalamus"
       167..544
       /codon_start=1
       /product="uncharacterized hypothalamus protein HBEX2"
       /protein_id="AA67654.1"
       /db_xref="GI:7689029"
       /translation="MESKEKRAVNSLSMENANOENKEQVANKGEPALPLADAGEYC
       VPRGMRPRVROPILQYRMDMRLGEPQARRREMERIGEVRQLMEKLERKQLS
       HSLRAVSTDPRPHDHDFCLMP"
BASE COUNT    256 a      170 c      221 g      181 t
ORIGIN
1
MetAlaSerLysValLysGlnValIleLeuAspLeuThrValGluLys.. 16
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
167 ATGGAGTCCAAAGAGAAAGAGACGAGTAACAGCTCAGCATGAAAAATGC 216
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 AspLysLysAsnLysLysGlyLysAlaSerLysGlnSerGluGluG 33
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
217 CAACCAAGAAATGAGAAAAAGAGCAAGTGTCTATTAAAGGGAGCCCT 266
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 IuSerHisHisLeuGluGluValGluAsnLysLysProGlyLysAsnVal 49
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
267 TGGCCCTCCCTTGTGATGCTGTGTAATCTGTGTGCTAGAGAAATCGT 316
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 ArgArg..LysValArgArgLeuValProAsnPheLeuTrpAlaIlePr 65
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
317 AGGGGCTTCCCGCTTAGCAGACCCATCCCTCATATAGATGGATATGAT 366
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
65 AsnArgHisValAsp.....HisSerGluGlyGluG 77
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 GCATAGAGCTTGGAGACCAACAGCAAGATGAGAGAGAGAAATATGAAA 416
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
77 IuValGlyArgPheValGlyValMetGluAlaLysArgHisSerLys 93
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
417 GGATTGGGAGAGAGCTGAGACAGCTGATGAAAAAGCTGAG..... 457
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
94 GluGlnGlnMetArgProTyrThrArgPheArgThrProGluPro.... 108
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
458 GAAAGACAGTTGAGTCACTAGTCGCGGCGAGTCAGCACTGACCCCTCA 507
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 ....AspAsnHisTyrAspPheCysLeuIlePro 118
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

OM of: US-09-327-750D-33 to: EST:* out_format : pfs
 Date: Mar 11, 2002 2:16 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
 -MODEL=frame+ p2n.model -DEV=xlh
 -O=/cgn2_1/USPTO_spool/US09327750/runat_11032002_101153_20308/app_query.fasta_1.1472
 -DB=EST -QWTF=fastap -SUFFIX=rst -GAOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.500 -GAPOF=4.500
 -GAPEXT=0.050 -XGAPOF=10.000 -XGAPEXT=0.500 -GAPOF=6.000
 -GAPEXT=7.000 -XGAPOF=10.000 -XGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blosun62 -TRANS=human40.cdi
 -LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=300
 -MINLEN=0 -MAXLEN=200000000 -USER=US09327750.@CGL1_1_5654
 -NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
 -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-327-750D-33
 Query length: 128
 Database: EST:
 Database sequences: 11351937
 Database length: 107921985
 Search time (sec): 4085.940000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_est1:AL022932	+	646.00	1047.10	551	AL022932 V8130k28 Beddington m
gb_est1:AW536404	-	646.00	1045.88	707	AW536404 G0104A11-3 NIA Mouse
gb_est1:BG064920	-	643.00	1040.25	642	BG064920 H3025D01-3 NIA Mouse
gb_est1:BG277659	+	633.00	1025.69	575	BG277659 ux47c11.y1 Soares.NMMA
gb_est1:BG6870503	+	633.00	1023.51	756	BG6870503 G02791422F1 NCI_CGAP.8
gb_est1:BG654459	-	615.00	997.65	499	BE654459 UI-M-AJ1-aha-f-10-0-UI
gb_est1:AW536974	-	599.00	970.56	577	AW536974 G0110H10-3 NIA Mouse
gb_est1:BF607762	+	554.00	895.17	585	BF607762 MYL 000704 Mouse 9-day
gb_est1:BG228077	-	553.00	895.87	587	BG228077 ux47c11.y1 Soares.NMMA
gb_est1:BF608209	+	516.00	835.09	650	BF608209 MYL 001088 Mouse 9-day
gb_est1:BE291071	+	505.50	818.94	583	BE291071 G0108G311F1 NCI_CGAP.4
gb_gss:AZ936393	-	493.50	798.78	637	AZ936393 ZM0193105F Mouse 10kb
gb_est1:BG669326	+	489.50	795.95	412	BG669326 DRNAEBF1 Rat DRG Libr
gb_est1:BG1183527	-	480.50	776.95	701	BG1183527 UNL-P-FN-by-f-07-0-UNL
gb_est1:AW536634	-	472.00	765.12	549	AW536634 G0106H08-3 NIA Mouse
gb_est1:BG655001	+	471.50	764.34	634	BG655001 ib44f06.y1 HR85 islet
gb_est1:BG472167	+	471.50	763.17	634	BG472167 G02273150F1 NIH_MGC.16
gb_est1:BG7971303	+	471.50	762.87	658	BG7971303 G02273150F1 NIH_MGC.84
gb_est1:BG714823	+	471.50	762.68	674	BG714823 G02677146F1 NIH_MGC.96
gb_est1:BG702285	+	471.50	762.49	690	BG702285 AV702285 ADB Homo sapi
gb_est1:AV702285	+	471.50	762.48	691	AV702285 AV702285 ADB Homo sapi
gb_est1:AL526209	+	471.50	762.43	695	AL526209 AL526209 LTI_NFL003_NE
gb_est1:AL526209	+	471.50	762.39	699	AL526209 AL526209 LTI_NFL003_NE
gb_est1:BG6820179	+	471.50	762.32	783	BE790774 G01581956F1 NIH_MGC.7
gb_est1:BG790774	+	471.50	761.48	794	AL526247 AL526247 LTI_NFL003_NE
gb_est1:AL526247	+	468.50	757.71	683	AL526247 AV702643 ADB Homo sapi
gb_est1:AV702643	+	468.50	757.71	683	AV702643 AV702643 ADB Homo sapi
gb_est1:BE314909	-	468.00	756.01	594	BE314909 G0114A0074F1 NIH_MGC.9
gb_est1:AL520931	-	468.00	756.51	717	AL520931 G01513201F1 NIH_MGC.71
gb_est1:AL520931	-	468.00	755.92	772	BE889313 G02674781F1 NIH_MGC.96
gb_est1:BE889313	-	468.00	755.54	810	BE889313 G02674781F1 NIH_MGC.96
gb_est1:BG709503	+	468.00	755.29	835	BG709503 G02782101F1 NIH_MGC.96
gb_est1:BG707398	+	468.00	755.29	835	BG707398 G02782101F1 NIH_MGC.96
gb_est1:BG709503	+	468.00	754.30	946	BG709503 G02380917F1 NIH_MGC.96
gb_est1:BG709503	+	468.00	754.30	946	BG709503 G02380917F1 NIH_MGC.96
gb_est1:BF038787	+	468.00	753.53	1042	BF038787 G01462113F1 NIH_MGC.6
gb_est1:BF038787	+	468.00	752.59	946	AL929703 au63f04.y1 Schneider
gb_est1:BG715659	+	464.50	750.65	734	BG715659 G02677263F1 NIH_MGC.96
gb_est1:BG715659	+	464.50	750.65	734	BG715659 G02677263F1 NIH_MGC.96
gb_est1:BG714974	-	461.00	745.07	726	BG714974 G02677263F1 NIH_MGC.96
gb_est1:BF967675	-	461.00	744.36	793	BF967675 G02287383F1 NIH_MGC.96
gb_est1:BG707734	-	461.00	744.12	817	BG707734 G02671255F1 NIH_MGC.96
gb_est1:BG707734	-	461.00	744.12	817	BG707734 G02671255F1 NIH_MGC.96
gb_est1:BG705843	+	457.00	737.68	813	BG705843 G02669329F1 NIH_MGC.96
gb_est1:BG705843	+	457.00	737.68	813	BG705843 G02669329F1 NIH_MGC.96
gb_est1:BG709325	+	456.50	736.72	828	BG709325 G02673345F1 NIH_MGC.96

gb_est1:AI291270 + 455.00 738.14 6.0e-32 511 | AI291270 qml6e06.x1 NCI_CGAP
 gb_est1:AL024066 + 453.00 734.67 9.3e-32 526 | AL024066 p8225a45 Beddington
 gb_est1:BE266012 + 453.00 733.49 1.1e-31 610 | BE266012 G01194731F1 NIH_MGC
 gb_est1:AI291126 + 453.00 732.75 1.2e-31 669 | AI291126 qml5f02.x1 NCI_CGAP
 gb_est1:BF237433 + 453.00 732.27 1.3e-31 711 | BF237433 G01842108F1 NIH_MGC

seq_name: gb_est1:AL022932

seq documentation block:
 LOCUS AL022932 551 bp mRNA EST 29-DEC-1999
 DEFINITION V8130k28 Beddington mouse dissected endoderm Mus musculus CDNA
 clone 528-3B10 5', mRNA sequence.
 ACCESSION AL022932
 VERSION AL022932.1 GI:6645505
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 551)
 Harrison, S.M., Dunwoodie, S.L., Arkell, R.M., Lehrach, H. and
 Beddington, R.S.
 Isolation of novel tissue-specific genes from cDNA libraries
 representing the individual tissue constituents of the gastrulating
 mouse embryo
 Development 121 (8), 2479-2489 (1995)
 95401865
 Contact: Wiles, M., Lehrach, H. and Avner, P.
 EEC Mouse Transcript Mapping Consortium
 Genoscope - CNS
 2, rue Gaston Cremieux, 91000 Evry, France
 Email: pavner@pasteur.fr
 Clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
 Berlin, Germany. Web site http://www.rzpd.de
 Seq primer: CGGCGCGAATTCCTCGGT
 High quality sequence only submitted.
 Vector: pSPoriT; Site: 1: NotI; site: 2: SalI;
 Cloned unidirectionally.
 Dissected endoderm 7.5 days.
 Average insert size: 1.2 kb (range: 0.2 - 2.kb).
 Location/Qualifiers
 1..551
 /organism="Mus musculus"
 /strain="C57Bl6 x DBA"
 /db_xref="taxon:10090"
 /clone_lib="Beddington mouse dissected endoderm"
 /tissue_type="dissected endoderm"
 /dev_stage="7.5 dpc"
 /note="Vector: pSPoriT; Site: 1: NotI; Site: 2: SalI; Cloned
 unidirectionally.~ High quality sequence only submitted.~
 Average insert size: 1.2 kb (range: 0.2 - 2.kb)"

BASE COUNT 153 a 124 c 191 g 83 t
 ORIGIN
 alignment_scores:
 Quality: 646.00 Length: 128
 Ratio: 5.168 Gaps: 0
 Percent Similarity: 97.656 Percent Identity: 92.969

alignment_block:
 US-09-327-750D-33 x AL022932

Align seg 1/1 to: AL022932 from: 1 to: 551

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
 142 ATGGAGTCCAAAGATCAAGCGTGAATAATCTCAACATGGAGATGACCA 191
 17 sglnLysLysGluGluLysGluGluLysProGlnAspThrLysArgG 34
 192 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 241

```

location/Qualifiers
1. 642
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /clone_name="G010A11"
   /clone_1kb="MIA Mouse E7.5 Embryonic Portion cDNA Library"
   /sex="unknown"
   /dev_stage="7.5dpc Embryo"
   /lab_host="DH10B"
   /note="Vector: pSPORT1 (Gibco/BRL Life Technology);
Site.1: SalI; Site.2: NotI; Total RNAs were extracted from
5 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an oligo(dT) primer [NotI primer-adaptor
(from GibcoBRL)]
5'-pGACTAGATTCAGATCGACAGCGCCGCTTTTTTT-3']
from 0.5µg of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by

```

seq_name:	gf_test2:BG064920
seq_documentation_block:	
LOCUS	BG064920 707 bp mRNA EST 26-JAN-2001
DEFINITION	H3025D01.3 NIA Mouse 15K CDNA Clone Set Mus musculus cDNA clone
	H3025D01 3', mRNA sequence.
ACCESSION	BG064920
VERSION	BG064920.1 GI:12547483
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 707)

AUTHORS Kargul G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, J.R.S., Carter, M.G. and Ko, M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT Other ESTs: H3025D01-5

Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://igsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3025 row: D column: 01
 Seq primer: -21M13 Forward
 High quality sequence stop: 707
 POLYA=Yes.

FEATURES

source
 1. 707
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3025D01"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dt)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
 Mol Genet 7: 1967-1978."
 BASE COUNT 154 a 209 c 146 g 198 t
 ORIGIN

alignment_scores:
 Quality: 643.00 Length: 128
 Ratio: 5.144 Gaps: 0
 Percent Similarity: 97.656 Percent Identity: 92.188
 alignment_block:
 US-09-327-750D-33 x BG064920/rev ..
 Align seg 1/1 to reverse of: BG064920 from: 1 to: 707
 1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspH1 17
 606 ATGGAGTCCAAAGATCAAGCGGTGAAATCTCAACATGGAGATGACCA 557
 17 sglnLysLysGlnGlyLysGlnGlyLysProGlnAspThrIleLysArgG 34
 556 TCAGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 507
 34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
 506 AGCCAGCTGTGCCCTGACCTCGAGGCTGGCAAAACTGTGCACCTAGA 457
 51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyArgTr 67

|||||
 456 GGAGTCCAGCGCGGTCCGGTTCGGCAGCCCATCGCTCATTATAGT 407
 67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
 406 GGACCTGATGACAGAGGTGGGGAGCCGCCAGGAAGGATGAGAGAGAGA 357
 84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
 356 ACGTACAGAGGTTTGGGGTGTATGTGAGACAGCTCATGTGAGAGAGCTGAGG 307
 101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProH1 117
 306 GAAAGGCAGCTGAGCCACAGCGTGGGGCGGTAGCACTGACCCGCTCA 257
 117 SHisAspHisAspGluPheCysLeuMetPro 128
 256 TCATGACCACCATGATGAGTTTGCCTCATGCC 223
 seq_name: gb_est2:BG277659

seq_documentation_block:

LOCUS BG277659 575 bp mRNA EST 21-FEB-2001
 DEFINITION ux47c11.v1 Soares_NMAX_maxillary.process Mus musculus cDNA clone
 IMAGE:3513237 5' similar to TR:Q9RLJ2 Q9RLJ2 BEX1 PROTEIN. ; mRNA
 sequence.
 ACCESSION BG277659
 VERSION BG277659.1 GI:13073183
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 575)
 NC-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: ux47c11.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1393813
 Seq primer: -40RP from Gibco
 High quality sequence stop: 462.
 Location/Qualifiers
 1..575
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="Soares_NMAX_maxillary.process"
 /tissue_type="maxillary process"
 /lab_host="DH10B (phage-resistant)"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
 was primed with a Not I - oligo(dt) primer [5',
 TGTACCAATCTGAAGTGGAGCGCGCGCTTTTTTTTTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 156 a 136 c 201 g 81 t
 ORIGIN

alignment_scores:
 Quality: 633.00 Length: 128
 Ratio: 5.105 Gaps: 0
 Percent Similarity: 96.875 Percent Identity: 91.406
 alignment_block:

Fax: 301 443 9890
Email: mEST@mail.nih.gov

CDNA Library Preparation: M.B. Soares Lab Clone Distribution: Researchers may obtain BMAP CDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements

Seq primer: M13 Reverse

FEATURES
source
Location/Qualifiers
1..499
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-N-AJI-aha-f-10-0-UI"
/clone_lib="NIH_BMAP_MOB_N"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NIH_BMAP_MOB_N library is a normalized library constructed from mouse olfactory bulbs. The tag is a string of 5 nucleotides present between the Not I site and the oligo-dT track. The library was constructed as described by Ronaldo, Lennon and Soares, Genome Research 6: 791-806, 1996. Tissue provided by Ms. Annie Novakovich, Zivic-Miller Laboratories."

BASE COUNT 146 a 108 c 179 g 66 t

alignment_scores:
Quality: 615.00 Length: 123
Ratio: 5.125 Gaps: 0
Percent Similarity: 97.561 Percent Identity: 92.683

alignment_block:

US-09-327-750D-33 x BE654459 ..

Align seg 1/1 to: BE654459 from: 1 to: 499

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHI 17
|||||
131 ATGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGGAGATGACCA 180
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArg 34
|||||
181 TCAGAAAAAGGAGGAGAGGAAGAAAGCCACAGATACCATCAGAGGG 230
34 luProValValalaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
231 AGCACCTGTGGCCCTGACCTCCGAGGTGGCAAAACTGFCACCTAGA 280
51 GlyGlyArgArgPheArgValArgGlnProLysSerHisTyrArgTr 67
|||||
281 GGAGTCCGAGCGGTTCGGGTTCCGAGGCCATCGCTCACTATAGATG 330
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluAla 84
|||||
331 GGACCTGATGACAGAGGTGGGACCCCGAGGAGGATGACAGAGAGA 380
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
|||||
381 ACGTACAGAGGTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 430
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHI 117
|||||
431 GAAGGCACCTGAGCCACAGCCTGCGGCGGTGTGTGTGTGTGTGTGT 480
117 sHisAspHisAspGlu 123
|||||
481 TCATGACCAACCATGATGAG 499

seq_name: gb_est1:AW536974

seq_documentation_block: 577 bp mRNA EST 31-AUG-2000
LOCUS AW536974
DEFINITION G0110H10-3 NTA Mouse E7.5 Embryonic Portion cDNA Library Mus musculus cDNA clone G0110H10 3', mRNA sequence.
ACCESSION AW536974
VERSION AW536974.1 GI:7179391
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 577)
AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraaja,R., Doi,H., Wood,W.H., III, Becker,K.G. and Ko,M.S.H.
TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
MEDLINE 20381348
COMMENT Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: G0110 row: H column: 10
Seq primer: -21M13 Forward
High quality sequence stop: 577
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..577
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G0110H10"
/clone_lib="NTA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site 1: SalI; Site 2: NotI. Total RNAs were extracted from 6 Embryo. The double-stranded cDNA was synthesized by Gibco's kit with an oligo(dT) primer [NotI primer-adaptor from GibcoBRL] [5'-pgACTAGTTCTAGATCGAGCGCGCCCTTTT-3'] from 0.5ug of mRNA. The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker li-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko."

BASE COUNT 137 a 160 c 126 g 154 t
ORIGIN
alignment_scores:
Quality: 599.00 Length: 118
Ratio: 5.209 Gaps: 0
Percent Similarity: 97.458 Percent Identity: 93.220
alignment_block:
US-09-327-750D-33 x AW536974/rev ..

us-09-327-750d-33.rst

Tue Mar 12 09:01:41 2002

```

Align seg 1/1 to reverse of: AW536974 from: 1 to: 577

11 LeuAsnMetGluAsnAspHisGlnLysGluGluLysGluLysGluLysPr 27
|||||
576 CTCACATGGAGATGACCATCAGAAAAGAGGAGAGAGAGAGAGAGAGCC 527
|||||
27 oGlnAspThrIleLysArgGluProValValAlaProThrPheGluAlaG 44
|||||
526 ACAAGATACCATCAGAAAGGAGCCAGCTGTGGCCCTGACCTCGAGGCTG 477
|||||
44 LysAsnCysAlaProArgGlyGlyArgArgPheArgValArgGln 60
|||||
476 GCAAAACTGTCCACTAGAGAGGTCGAGGCGGTTCGCGGTTCGCGAG 427
|||||
61 ProIleSerHisTyrArgTrpAspLeuMetHisArgValGlyGluProG 77
|||||
426 CCCATCGCTCATATAGTAGGAGCTGATGAGAGGCTTGGGAGGCCCA 377
|||||
77 nGlyArgMetArgGluGluAsnValGlnArgPheGlyGluAspMetArg 94
|||||
376 GGAAGGATGAGAGAGGAGACCTACAGAGCTTGGGGGTGATGTGAGAC 327
|||||
94 InLeuMetGluLysLeuArgGluArgGlnLeuSerHisSerLeuArgAla 110
|||||
326 AGCTCATGAGAGAGCTGAGGAGAGGACCTGACGACGACGACGCTGCG 277
|||||
111 ValSerThrAspProHisHisAspHisHisAspGluPheCysLeuMe 127
|||||
276 GTTAGCACTGACCGGCTCATGACCATGATGAGTTTGGCTCAT 227
|||||
127 tPro 128
|||||
226 GCCC 223

seq_name: gb_est2:BF607762

seq_documentation_block: 785 bp mRNA EST 01-APR-2001
LOCUS BF607762
DEFINITION M1.000704 Mouse 9-day fetus cDNA library ICRFp522 Mus musculus
cDNA clone ICRFp522A1947 5', mRNA sequence.
ACCESSION BF607762
VERSION BF607762.1 GI:13504254
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 785)
AUTHORS Yahyawi,M., Hennig,S., Neidhardt,L., Radelof,U., Hermann,B.G.,
Lehrach,H. and O'Brien,J.
TITLE Detection of a high number of novel genes in a 9-day mouse embryo
cDNA library normalised by oligonucleotide fingerprinting
JOURNAL Unpublished (2001)
COMMENT Contact: Hennig S
Laboratory 123, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inhnstr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennig@molgen.mpg.de
EST's are made from clones being representatives of clone clusters.
Clone clusters were calculated from oligonucleotide fingerprints.
PCR Primers
FORWARD: 5'-GAGCTATTCAGAGTAGTA-3'
BACKWARD: 5'-TAATACGACCTACTATAGG-3'
Seq primer: 5'-ATTAGGTGACACTATAG-3'
High quality sequence stop: 785.
FEATURES
Location/Qualifiers
1..785
source
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="ICRFp522A1947"
/clone_lib="Mouse 9-day fetus cDNA library ICRFp522"

/tissue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E.coli XL1 blue"
/notes="Vector: pSVSPORT1; Site_1: NotI; Site_2: SalI;
Library preparation by oligo-dr priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."
BASE COUNT 190 a 188 c 233 g 153 t 21 others
ORIGIN

alignment_scores:
Quality: 554.00 Length: 128
Ratio: 4.617 Gaps: 4
Percent Similarity: 93.750 Percent Identity: 87.500

alignment_block:
US-09-327-750D-33 x BF607762 ..
Align seg 1/1 to: BF607762 from: 1 to: 785

1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHi 17
259 ATGGAGTCCAAAGATCAAGGCGTGAATAATCTCAACATGGAGATGACCA 308
|||||
17 sGlnLysLysGluGluLysGluLysProGlnAspThrIleLysArgG 34
|||||
309 TCAGAAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 358
|||||
34 LuProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
|||||
359 AGCCAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAACTGTGCACCTAGA 408
|||||
51 GlyGlyArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
|||||
409 GGAGGTGCGAGGCGGTTCGGGTTCGGCAGGCCATCGCTACTATAGATG 458
|||||
67 PAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGlu 84
|||||
459 GGACCTGTATGCAGAGGTTGGGAGGCCGCCAGGAGGATGAGAGAGAGA 508
|||||
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMet.GluLysLeuAr 100
|||||
509 ACGTACAGAGGTTTGGGGGTGATGTGAGACAGCTCATGGGGGAGCTGAG 558
|||||
100 gGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAsp.ProPro 116
|||||
559 GGAAGGCACTCAACACAGAGCTGCGGGCGGTAGCAGCTGACCCCGCCT 608
|||||
117 His.HisAspHisHis.AspGluPhe 124
|||||
609 CATTTCATGACCACCATGGATGAGTTT 634

seq_name: gb_est2:BG228077

seq_documentation_block: 587 bp mRNA EST 08-FEB-2001
LOCUS BG228077
DEFINITION ux47c11.x1 Soares.NMXX maxillary process Mus musculus cDNA clone
IMAGE:3513237 3' similar to TR:Q9RIJ2 Q9RIJ2 BEX1 PROTEIN. ; mRNA
sequence.
ACCESSION BG228077
VERSION BG228077.1 GI:12715592
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 587)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
TITLE NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

```


Tue Mar 12 09:01:41 2002

```

456 ACGTACAGAGCTGGGGTGTGTGACAGAGCTCATGGAGAACTCAGG 505
101 GluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspProProHis 117
506 GAAAGGAGCTTTAGCCACAGCTGGGGGCTTACGACTGACGCGCTCA 555
117 sHisAspHisAspGluPheCysLeuMetPro 128
556 TCATGACTACCATGATTAGTTTGGCTCATGCC 589

seq_name: gb_estl:BE291071
seq_documentation_block: 583 bp mRNA EST 13-JUL-2000
LOCUS BE291071
DEFINITION 601086311F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3500522 5',
mRNA sequence.
ACCESSION BE291071
VERSION BE291071.1 GI:9172545
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 583)
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM8559 row: e column: 03
High quality sequence start: 5
High quality sequence stop: 503.
FEATURES
Location/Qualifiers
1..583
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="3500522"
/clone_lib="NCI_CGAP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt:
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies, Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 165 a 136 c 201 g 81 t
ORIGIN

alignment_scores:
Quality: 505.50 Length: 120
Ratio: 4.554 Gaps: 2
Percent Similarity: 92.500 Percent Identity: 83.333

alignment_block:
US-09-327-750d-33 x BE291071
Align seg 1/1 to: BE291071 from: 1 to: 583
1 MetGluSerLysAspGlnGlyAlaLysAsnLeuAsnMetGluAsnAspHis 17
208 ATGGAGTCCAAATGACAGCGGTGAAATCTCACATGGAGATGACCA 257
17 sGlnLysLysGluGluLysGluGluLysProGlnAspThrIleLysArg 34

```

```

258 TCAGAAAGAGGAGAGAGGAGAAAGCCACCAAGATACCATCAGAAGG 307
34 luProValValAlaProThrPheGluAlaGlyLysAsnCysAlaProArg 50
308 ACCAGAGCTGTGGCCCTGACCTCCGAGGCTGGCAAAAACTGTGCACCTAGA 357
51 GlyGlyArgArgArgPheArgValArgGlnProIleSerHisTyrArgTr 67
358 GGAGGTCGAGCGGCTTCGGGTTCCGCGAGCCCATCGCTCACTATAGATG 407
67 pAspLeuMetHisArgValGlyGluProGlnGlyArgMetArgGluGluA 84
408 GGACCTGATGACAGAGGTTGGGGAGCCCACTGAGGATGAGAGAGGAGA 457
84 snValGlnArgPheGlyGluAspMetArgGlnLeuMetGluLysLeuArg 100
458 ACCTACAGAGGTTAGGGGGTGATGTGAGACAGCTCATGGAGAAGCTGAGG 507
101 GluArgGlnLeuSerHis...SerLeuArgAlaValSerThrAspProp 116
508 GAAAGGCTGTGGAGCCCAAGACATCGGAGCGGTTAGACATGACCCGA 557
116 roHishis 118
558 CTCATCAT 565

seq_name: gb_gss:AZ936393
seq_documentation_block: 637 bp DNA GSS 26-APR-2001
LOCUS AZ936393
DEFINITION 2M0193105F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0193L05 F, DNA sequence.
ACCESSION AZ936393
VERSION AZ936393.1 GI:13794974
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 637)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0193 row: L column: 05
Seq primer: CGTTGTAACACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 637.
FEATURES
Location/Qualifiers
1..637
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="UUGC2M0193L05"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pWDA2nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource

```

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The ligated DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (9147321419b1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 131 a 197 c 132 g 176 t 1 others
ORIGIN

alignment_scores: Quality: 493.50 Length: 117
Ratio: 4.569 Gaps: 1
Percent Similarity: 92.308 Percent Identity: 84.615

alignment_block:

US-09-327-750D-33 x AZ936393/rev ..

Align seg 1/1 to reverse of: AZ936393 from: 1 to: 637

```

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
|||||.....:|||||.....:|||||.....:|||||.....:
302 ATGAGATGCCAAGTGGACACAGCGCTGAAATAATCTCAACATGAGAGATGA 333
16 PHisGlnLysGlnGluLysGlnGluLysProGlnAspThrIleLysA 33
|||||.....:|||||.....:|||||.....:|||||.....:
332 CCATCAGAGAAAGGAGAAAGGAGAAAGGAGAGAGAGAGAGAGAGAGAG 283
33 TGGTProValValAlaProThrPheGluAlaGlyLysAsnGlyAlaPro 49
|||||.....:|||||.....:|||||.....:|||||.....:
282 GGGATCCGATGTTGGCCCTCTTCGAAAGCTGCACTACTACATGCTGCT 233
50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||.....:|||||.....:|||||.....:|||||.....:
232 AGAGAGAGTCCAGCGGCTTCGGGCTTCGGCAGCCATGTCACATCAG 183
66 GTPAspleuMethHisArgValGlyLysProGlnGlyArgMetArgGlu 83
|||||.....:|||||.....:|||||.....:|||||.....:
182 ATGGAGCTGATGATAGGTTGGGAGCCCCAGGAGAGAGATGAGAGAG 133
83 LAsnValGlnArgPheGlyLysAspMetArgGlnLeuMetGluLysLeu 99
|||||.....:|||||.....:|||||.....:|||||.....:
132 AGAAGCTACAGAGGTTGGGATGATGTGAGACAGCTCAGAGAGAGCTN 83
100 ArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrAspPro 116
|||||.....:|||||.....:|||||.....:|||||.....:
82 GG-GAAAGGACAGTGAAGCCAGCCTCGGGCGGTTAGCAGTACCCGCC 34
116 o 116
1
33 T 33

```

seq_name: gb_est2:BG669326

seq_documentation_block:

LOCUS BG669326 412 bp mRNA EST 30-APR-2001
DEFINITION DRNAEB11 Rat DRG Library Rattus norvegicus cdna clone DRNAEB11
5' mRNA sequence.
ACCESSION BG669326
VERSION BG669326.1 GI:13891248
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 412)
AUTHORS Xiao H.S., Han Z.G., Zhang F.X., Huang Q.H., Lu Y.J., Bao L., Fu G., Guo C., Yan Q., Jin S.X., Zhu Z.D., Xu X.R., Li N.G., Chen Z. and Zhang X.
TITLE Distinct gene expression profiles of rat dorsal root ganglion induced by peripheral nerve axotomy
JOURNAL Unpublished (2001)
COMMENT Contact: Zhang Xu
Laboratory of Sensory System
Institute of Neuroscience
320 Yue Yang Road, Shanghai 200031, P.R.China
Tel: 86-21-64748700-121
Fax: 86-21-64713446
Email: xu.zhang@ion.ac.cn

FEATURES
source Location/Qualifiers
1..412
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="DRNAEB11"
/clone_lib="Rat DRG Library"
/sex="male"
/tissue-type="dorsal root ganglion"
/dev-stage="adult"
BASE COUNT 124 a 83 c 130 g 75 t
ORIGIN

alignment_scores: Quality: 489.50 Length: 126
Ratio: 4.450 Gaps: 2
Percent Similarity: 87.302 Percent Identity: 76.984

alignment_block:

US-09-327-750D-33 x BG669326 ..

Align seg 1/1 to: BG669326 from: 1 to: 412

```

1 MetGluSerLys...AspGlnGlyAlaLysAsnLeuAsnMetGluAsnAs 16
|||||.....:|||||.....:|||||.....:|||||.....:
37 ATGGATCCGATGTTGGCCCTCTTCGAAAGCTGCACTACTACATGAGAA 86
16 PHisGlnLysGlnGluLysGlnGluLysProGlnAspThrIleLysA 33
|||||.....:|||||.....:|||||.....:|||||.....:
87 CCATCAGAGAAAGGAGAAAGGAGAAAGGAGAGAGAGAGAGAGAGAGAG 136
33 TGGTProValValAlaProThrPheGluAlaGlyLysAsnGlyAlaPro 49
|||||.....:|||||.....:|||||.....:|||||.....:
137 GGGAGCGGTTGTAGCCCTGCTTCGAAAGCTGAGAAATACATGCTGCT 186
50 ArgGlyArgArgArgPheArgValArgGlnProIleSerHisTyrAr 66
|||||.....:|||||.....:|||||.....:|||||.....:
187 ATAGGAAGTCCAGCGGTTCCGCTTCGGCAGCCATTCCTCCTACTATAT 236
66 g.TTPAspleuMethHisArgValGlyLysProGlnGlyArgMetArgGlu 82
|||||.....:|||||.....:|||||.....:|||||.....:
237 GATGGAGCTGATGCACAAGGTTGAGAGAGCCCAAGGAGAGATGAGAGA 286
83 GlnAsnValGlnArgPheGlyLysAspMetArgGlnLeuMetGluLysLe 99
|||||.....:|||||.....:|||||.....:|||||.....:
287 GAGAATGTCAAGAGTTGGGAGAGATATGAGACAGCTCATGTGTAAT 336

```



```

/db_xref="taxon:10090"
/clone_1lb="NIA Mouse E7.5 Embryonic Portion cDNA Library"
/sex="unknown"
/dev_stage="7.5dpc Embryo"
/lab_host="DH10B"
/notes="vector: pSPORT1 (Gibco/BRL life technology);
site_1: SalI; site_2: NotI. Total RNAs were extracted from
6 Embryo. The double-stranded cDNA was synthesized by
Gibco's kit with an Oligo(dt) primer [NotI primer-adaptor
from GibcoBRL]
[5'-pGACTAGTTCAGATCCGAGCGCCGCCCTTTTCTTTT-3']
from 0.5ug of mRNA. The double-stranded cDNAs were
treated with T4 DNA polymerase and purified by
ethanol-precipitation. The cDNAs were ligated to
lone-linker lI-SalI (include SalI sequence). The cDNAs
were purified by phenol/chloroform and separated from
free linkers by Centricon 100. Then, cDNAs were amplified
by long-range high fidelity PCR using Takara's Ex Taq
polymerase. Then, the cDNAs were purified by
phenol/chloroform and by Centricon 100. The cDNAs were
digested with SalI and NotI enzymes. Then, the cDNAs were
size selected by Gibco's Size Fractionation Column. The
cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by chemical method. The library was
constructed by Xiaohong Wang and Minoru S. H. Ko."

```

BASE COUNT
ORIGIN

151 a 151 c 117 g 130 t

alignment_scores:
Quality: 472.00 Length: 97
Ratio: 5.187 Gaps: 0
Percent Similarity: 93.814 Percent Identity: 88.660

alignment_block:
US-09-327-750D-33 x AWS36634/rev ..

Align seg 1/1 to reverse of: AWS36634 from: 1 to: 549

```

32 LysArgGluProValAlaProThrpheGluAlaGlyLysAncysAl 48
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
541 AAAAGGATCCGATTGTGGCCCTGCTTCAAGCTGAGACACTACTACT 492
48 aProArgGlyGlyArgArgArgPheArgValArgGlnProIleSerHisT 65
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
491 GCCTAGAGAGAGGTCCGAGCGGTTCGGGTTCGGCAGCCCATCTGTGCACT 442
65 yTrArgTrpAspLeuMetHisArgValGlyGluProGlnGlyArgMetArg 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
441 ACGATGGGACCTGATGCATAGGTGGGAGCCCGAGGAGAGATGAGA 392
82 GluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetGluIly 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
391 GAGGAGAAAGTACAGAGGTTGGGATGATGTGAGACAGCTCATGGAGAA 342
98 sLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThrasp 115
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
341 GCTGAGGAGAAAGCAAGCTGAGCCACAGCTGGGGCGGTATGACACTGACC 292
115 roProHisHisAspHisAspGluPheCysLeuMetPro 128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
291 CGCCTCATCATGACCACCATGATGAGTTTGGCTTATGCCC 251

```


Tue Mar 12 09:01:40 2002

us-09-327-750d-33.rml

Page 1

OM of: US-09-327-750D-33 to: Issued_Patents_NA.* out_format: pfs
Date: Mar 11, 2002 3:51 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters: -DEV-X1h
-MODEL=frame+g2n.model -DEV-X1h
-O=/cgn2_1/USPTO.spool/US09327750/r/unatc_11032002_101154_20340/app_query.fasta.1.1472
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rm -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -GAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORMEXT=HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09327750 -CGN1_1.99 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPEX -WAIT -THREADS=1

Search information block:
Query: US-09-327-750D-33
Query length: 128
Database: Issued_Patents_NA.*
Database sequences: 351203
Database length: 113238999
Search time (sec): 146.090000

Score list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-381-881-1 +		93.50	166.92	0.1200	1315
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-281-221-1 +		93.50	166.92	0.1200	1315
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-07-951-715A-6 +		93.00	155.63	0.5107	3624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-459-448A-6 +		93.00	155.63	0.5107	3624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-459-504B-6 +		93.00	155.63	0.5107	3624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-459-504B-6 +		93.00	155.63	0.5107	3624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-459-444-6 +		93.00	155.63	0.5107	3624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-053-549-1 +		91.00	142.54	2.74	8854
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-506-553C-6 +		82.50	152.65	0.7483	607
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-998-416-994 -		82.00	149.79	1.08	729
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-426-599B-3 +		81.00	149.35	1.14	624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-147-777-3 +		80.50	110.90	158.26	24979
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-452-872-3 +		80.50	110.90	158.26	24979
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-434-288-10 +		80.00	137.79	5.03	1600
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-030-995-1 +		79.00	137.39	5.30	1365
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-117-927-4 +		78.50	131.71	10.97	2163
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-931-999-4 +		77.50	118.14	62.53	6755
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-125-468-1 -		77.50	103.01	435.42	30001
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-474-933-1 -		77.00	140.56	3.53	672
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-998-416-141 -		77.00	133.96	8.23	1288
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-047-041A-24 -		76.50	130.76	12.40	1599
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-630-349-11 +		76.50	127.81	18.09	2137
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-285-440-13 +		76.50	127.81	18.09	2137
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-330-349-13 +		76.50	127.81	18.09	2137
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-306-691B-23 +		76.50	127.06	19.92	2301
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-056-200-93 +		76.50	112.62	127.02	9551
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-800-644-93 +		76.50	112.62	127.02	9551
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-036-967A-1 -		76.00	90.02	2.3e+03	80161
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-370-700-1 -		76.00	90.02	2.3e+03	80161
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-717-294-42 +		75.50	118.36	60.85	4451
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-103-840A-2 +		75.00	137.29	5.37	624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-381-881-1 +		75.00	137.29	5.37	624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-09-281-221-1 +		75.00	137.29	5.37	624
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-809-740A-1 -		75.00	120.54	46.01	3252
/cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-822-586-45 +		75.00	109.04	201.00	10095

seq_name: /cgn2_6/prodata/2/ina/5B.COMB.seq:US-08-381-881-1

seq_documentation_block:

Sequence 1, Application US/08381881
Patent No. 593064

GENERAL INFORMATION:
APPLICANT: SAYELKOU, PAUL H. M.
INVENTOR: GASTRICK, WILEM
TITLE OF INVENTION: BORDETELLA BRONCHISEPTICA VACCINE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
STREET: 515 NORTH WASHINGTON STREET.
CITY: ALEXANDRIA
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/381,881
FILING DATE:
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: POULOS III, JAMES A.
REGISTRATION NUMBER: 31,714
REFERENCE/DOCKET NUMBER: TTP 29685
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703 549-7200
TELEFAX: 703 528-5313

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1315 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bordetella bronchiseptica
STRAIN: 401
IMMEDIATE SOURCE:
CLONE: E coli PC2495(pivB3-420)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..539
FEATURE:
NAME/KEY: CDS
LOCATION: 540..1142
FEATURE:
NAME/KEY: misc feature
LOCATION: 1143..1315

US-08-381-881-1

alignment_scores:
Quality: 93.50 Length: 96
Ratio: 1.948 Gaps: 4
Percent Similarity: 50.000 Percent Identity: 33.333

alignment_block:
US-09-327-750D-33 x US-08-381-881-1

```
Align seg 1/1 to: US-08-381-881-1 from: 1 to: 1315
43 AlaGlyLysAsnCysAlaProArgGlyArgArgPheArgValar 59
|||||:||||| ||| ||||| |||||:|||||
542 GCAAGCCAAACAGTTCCTCTGGCGCGCGCTCGCGCGCTCG 591
59 gGlnProle.....SerHisTyrArgTrrp..... 67
|||||:||||| ||||| ||||| |||||
592 CGGCCCATGCCGAGACGCCACCATTCATTACCGGCACGATCACCGAC 641
68 ..AspLeuMetHisArgValGlyGluPro.....GlnGlyArgMet 80
|||||:||||| ||| ||||| ||||| |||||
642 CAGACCTGCAGATCGAGGACCGCGCTTACATCAAGGTCTGCA 691
81 ArgGluGluAsnValGlnArgPheGlyGluAspMetArgGlnLeuMetG1 97
|||||:||||| ||||| ||||| |||||
692 CCT.....GCCACGATCTCCAAGAGCGCGCTGA 720
97 uLysLeuArgGluArgGlnLeuSerHisSerLeuArgAlaValSerThra 114
|||||:||||| ||||| ||||| ||||| |||||
721 AGAAGCGCGCGAGCTGGCGGGCGCACCTCGCTTCGATATCAAGCTGAAG 770
114 spProProHisHisAspHisArgPheCysLeu 126
|||||:||||| ||||| ||||| |||||
771 GACTCGCGCGACCGCTCAACACTCTCAAGCTGTACTT 808

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-281-221-1
seq_documentation_block:
; Sequence 1, Application US/09281221
; Patent No. 6284256
; GENERAL INFORMATION:
; APPLICANT: SAEVLAOUL, PAUL H., M.
; TITLE OF INVENTION: BORDETTELLA BRONCHISEPTICA VACCINE
; NUMBER OF SEQUENCES: 6
; ADDRESSEE: STEVENS, DAVIS, MILLER & MOSHER
; STREET: 515 NORTH WASHINGTON STREET.
; CITY: ALEXANDRIA
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 30-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/381,881
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: POULOS III, JAMES A.
; REGISTRATION NUMBER: 31,714
; REFERENCE/DOCKET NUMBER: TPP 29685
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 549-7200
; TELEFAX: 703 528-5313
; TELEX: 44-0704 "STEVENS"
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

Page 3

[illegible]

```

; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; MOLECULE TYPE: OTHER nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; OTHER INFORMATION: optimized cryiB"
; OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-08-459-448A-6

alignment_scores:
Quality: 93.00 Length: 116
Ratio: 1.550 Gaps: 4
Percent Similarity: 51.724 Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-08-459-448A-6 ..
Align seg 1/1 to: US-08-459-448A-6 from: 1 to: 3624

14 GluAsnAspHisGlnLysGluLysGluLysGluLysProGlnAspTh 30
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1026 GAACACCAGCACCAGCGGCCACCAACCCAGCATCAACCCGTGACCC 1075

30 rleLysArgGluProValValalaProThrPheGluAlaGlyLysAsnC 47
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1076 TGGCGCTTCGCCAGCGCGAGCGTGTACCGCACCGAGAGACTACCCGGCGTG 1125

47 ysAla.....ProArgGlyGlyArgArg 55
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1126 CTGCTGGGGCATCTACCTGGAGCCCATCCAGCGGTGCCACCGTGC 1175

56 PheArgValArgLnpProIleSerHisTyraGTrpAspLeuMetHisar 72
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1176 CTTCACTTCACCAACCCCCAGACATCACGGA.....CCGCGCACCG 1219

72 gvalGlyGluProGInglyArGmetArGgLuGluAnValGInArGPheG 89
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1220 CCAACTACAGCCA.....GCCCTACGAGAGCCCGCGCTGCAGCTGAAG 1263

89 lyGluAspMetArgGlnLeuMetGluLysLeuArgGluArgGlnLeuSer 105
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1264 GACAGCCAGACGAGCTCCCCCGGACACCAACCCAGCGGCCCAA..... 1307

106 HisSerLeuArgAlaValserThrAspProHisHisaspHis 121
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1308 .....CTAGGAGCTACGACCCCGCTGAGCCACATCGGCATCAT 1349

seq_name: /cgn_6/ptodata/2/ina/6A_COMB.seq:US-08-459-595A-6
seq_documentation_block:
Sequence 6. Application US/08459595A
Patent No. 6018104
GENERAL INFORMATION:
APPLICANT: Koziel, Michael G.
APPLICANT: Desai, Nallini M.
APPLICANT: Lewis, Kelly S.
APPLICANT: Kramer, Vance C.
APPLICANT: Warren, Gregory W.
APPLICANT: Evola, Stephen V.
APPLICANT: Crossland, Lyle D.
APPLICANT: Wright, Martha S.
APPLICANT: Merlin, Ellis J.
APPLICANT: Launis, Karen L.
APPLICANT: Rothstein, Steven J.
APPLICANT: Bowman, Cindy G.
APPLICANT: Dawson, John L.
APPLICANT: Dunder, Erik M.
APPLICANT: Pace, Gary M.
APPLICANT: Sutlie, Janet L.

```

```

56 PheArgValArgInProIleSerHisTyrArgTyrAspLeuMetHisAr 72
   ::::::::::::::::::::
1176 CTTCACCTTACACACCCGAGACATCAGCA.....CCCGGCGCCG 1219
      72 gvalgluprogInglYArgmetArgIugluasnValGlnArgPheg 89
      1220 CCACACTACAGCA.....GCCCTACGACAGCCCGCGCTGAGCTGAG 1263
      89 lyluaspmetArgInLeuMetArgIulysLeuArgIulnArgInleuser 105
      1264 GACAGCGAGACGAGCTGCCCCCGAGACACCGAGCGCCCA..... 1307
      106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
      1308 .....CTACGAGAGCTACAGCCGCGCTGAGCCACATCGGCATCAT 1349

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-459-504B-6

seq_documentation_block:
; Sequence 6, Application US/08459504B
; Patent No. 6075185
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; APPLICANT: Rothenstein, Steven J.
; APPLICANT: Bowman, Cindy G.
; APPLICANT: Dawson, John L.
; APPLICANT: Dunder, Erik M.
; APPLICANT: Pace, Gary M.
; APPLICANT: Suttie, Janet L.
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
; TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 6075185artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459, 504B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/459, 595
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 07/951, 715
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/772, 027
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 3624 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Synthetic DNA"
HYPOTHEICAL: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3621
OTHER INFORMATION: /product= "Full-length, maize
                    optimized cyr1B
                    OTHER INFORMATION: /note= "Disclosed in Figure 6."
US-08-459-504B-6

alignment_scores:
Quality: 93.00 Length: 116
Ratio: 1.550 Gaps: 4
Percent Similarity: 51.724 Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-08-459-504B-6 ..
Align seg 1/1 to: US-08-459-504B-6 from: 1 to: 3624

14 GluAsnAspHisGlnIlylYsGluGluGluLysProGlnAspH 30
   |||:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1026 GAACACCCAGACCCAGCGGCGCCACCAACACAGCATCAACCCGCTAGCC 1075
      30 rlllysArgIuproValAlaIapProThrPheGluAlaGlyLysAsnC 47
      1076 TCGCGTCCGCGACCGCGGACGTGTACCGCACGAGAGCTTACGGCGCGTG 1125
      47 ysAla.....ProArgIlylArgArgArg 55
      1126 CTGCTGTGGGCGATCTACTGAGCCATCCACGGCGGTGCCACCGTGGG 1175
      56 PheArgValArgInProIleSerHisTyrArgTyrAspLeuMetHisAr 72
      1176 CTTCACCTTACACACCCGAGACATCAGCA.....CCGCGGACCG 1219
      72 gvalgluprogInglYArgmetArgIugluasnValGlnArgPheg 89
      1220 CCACACTACAGCA.....GCCCTACGAGACCCCGCGCTGAGCTGAG 1263
      89 lyluaspmetArgInLeuMetArgIulysLeuArgIulnArgInleuser 105
      1264 GACAGCGAGACGAGCTGCCCCCGAGACACCGAGCGCCCA..... 1307
      106 HisSerLeuArgAlaValSerThrAspProHisHisAspHisHis 121
      1308 .....CTACGAGAGCTACAGCCGCGCTGAGCCACATCGGCATCAT 1349

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-459-444-6

seq_documentation_block:
; Sequence 6, Application US/08459444A
; Patent No. 6121014
; GENERAL INFORMATION:
; APPLICANT: Koziel, Michael G.
; APPLICANT: Desai, Nalini M.
; APPLICANT: Lewis, Kelly S.
; APPLICANT: Kramer, Vance C.
; APPLICANT: Warren, Gregory W.
; APPLICANT: Evola, Stephen V.
; APPLICANT: Crossland, Lyle D.
; APPLICANT: Wright, Martha S.
; APPLICANT: Merlin, Ellis J.
; APPLICANT: Launis, Karen L.
; TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED
; NUCLEIC ACID CODING SEQUENCE

```

```

alignment_scores:
  Quality: 93.00      Length: 116
  Ratio: 1.550       Gaps: 4
  Percent Similarity: 51.724  Percent Identity: 26.724

alignment_block:
US-09-327-750D-33 x US-08-459-444-6 ..

Align seg 1/1 to: US-08-459-444-6 from: 1 to: 3624

14  GluAsnAspHisGlnLysLYSLysGluGluLysGluGluLysProGlnAspH 30
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1026 GAACACACGACCCACCGCGCCACCAACACGACGATCAACCCCGTGACCC 1075
                                     |||:|||||:|||||:|||||
30  rleLysAsrGluProValAlaIaProThrPheGluAlaGlyLysAsnC 47
                                     |||:|||||:|||||:|||||
1076 TGGCGCTTCGCCGCGCGACGTGTACCGACCGAGAGCTACGCCGGCGTG 1125

47  ysala.....ProArGGLyArGArGArg 55
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1126 CTCTCTGGGGCATCTACTGGAGCCCATCAACGCGGTGCCACCGCGCG 1175

56  PheArgValArgGlnProIleSerHisTyrArgTirPaspLeuMetHisAr 72
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1176 CTTCACACTTCACACCCCGCAACATCAGGGA.....CGCGGGCGACCG 1219

```

```

alignment_scores:
  Quality: 93.00      Length: 116
  Ratio: 1.50         Gaps: 4
  Percent Similarity: 51.724  Percent Identity: 26.724

alignment_block:
  US-09-327-750D-33 x US-09-053-549-7 ..
  Align seg 1/1 to: US-09-053-549-7 from: 1 to: 3624

```

Align seq 1/1 to: US-09-053-549-7 from: 1 to: 3624